The Palgrave Handbook of Deceptive Communication

Edited by
Tony Docan-Morgan

The Palgrave Handbook of Deceptive Communication

Tony Docan-Morgan
Editor

The Palgrave Handbook of Deceptive Communication

Editor
Tony Docan-Morgan
Department of Communication Studies
University of Wisconsin-La Crosse
La Crosse, WI, USA

Cover image: © Mike Hill

PREFACE

The Palgrave Handbook of Deceptive Communication unravels the topic of lying and deception in human communication, offering an interdisciplinary and comprehensive examination of the field, rethinking current approaches to the subject, presenting original research, and offering direction for future investigation and application. Scholars from around the world investigate historical perspectives on the study of lying and deception, the myriad forms of deceptive behavior, cross-cultural perspectives on deceit, moral dimensions of deceptive communication, theoretical approaches to the study of deception, and strategies for detecting and deterring deceit. Truth-telling, lies, and the many gray areas in-between are explored in the contexts of identity formation, interpersonal relationships, groups and organizations, social and mass media, marketing, advertising, law enforcement interrogations, court, politics, and propaganda. This handbook is designed for advanced undergraduate and graduate students, academics, researchers, practitioners, and anyone interested in the pervasive nature of truth, deception, and ethics in the modern world.

The Palgrave Handbook of Deceptive Communication is particularly unique because of its diverse disciplinary, methodological, theoretical, and applied perspectives. More than 100 prominent and emerging deception scholars have generously written 51 chapters for this volume. Contributors come from an array of fields, including communication studies, psychology, sociology, anthropology, philosophy, ethics, law, criminology and forensic science, psychiatry and behavioral neuroscience, counseling, literature, linguistics, business, management, journalism, advertising, public relations, marketing, and political science. The substantial collection of multidisciplinary knowledge in the current handbook serves as platform from which new questions, investigations, and discoveries will emerge.

The term *deceptive communication* was chosen purposefully for the title of this handbook. Stated simply, *human deception is communicative.* It involves senders and receivers, information or message exchange, and consequences that can be small or large, short- or long-term, personal or public. What makes deception distinct from many other forms of communication is that at least one communicator in the interaction—regardless of context—manipulates words, behaviors, texts, objects, and/or appearances so that others will form a false impression. Stated more directly, deceptive communication is the exchange of information that is known by the communicator to be inaccurate and/or misleading. What constitutes deception in daily interaction is context dependent and based on communicators' perceptions of various factors such as intent and awareness. When we approach the topic of deception as a multifaceted communicative phenomenon, we are better able to explain, describe, predict, and in some instances control how deception and truth-telling unfold.

The communication of truth and deception is a phenomenon that has affected and will continue to affect humans in their personal, professional, and civic lives. Yet, the ability to ask questions and find answers about deceptive communication is a powerful force. As we search for and find truth about deceptive communication, the realized and potential outcomes are substantial—becoming more literate consumers of information, more skilled message creators, and more impactful in creating a just and humane world.

La Crosse, WI, USA   Tony Docan-Morgan

Acknowledgements

This handbook moved from an idea to reality with the support of a number of individuals and organizations, for which I am extremely grateful.

First, thank you to all of the outstanding contributors for your hard work, deep insight, and commitment to uncovering truth. I hope our collaboration in this volume leads to future endeavors with one another. I express my deepest gratitude to all of you.

Second, I am grateful to the University of Wisconsin-La Crosse (UWL) for providing support in the form of a sabbatical to work on this volume. I am thankful to UWL's excellent leadership, as well as the institution's commitment to fostering curiosity and life-long learning through collaboration, innovation, and the discovery and dissemination of new knowledge. I am also grateful to be a part of UWL's College of Arts, Social Sciences, and Humanities, where the mission of fostering intellectual curiosity and creativity is a reality. Thank you specifically to Dr. Linda Dickmeyer for your support and encouragement, my colleagues in the Department of Communication Studies and throughout UWL, and the many curious and engaged students I have had over the years. I am proud to call UWL my academic home.

Third, I offer my gratitude to Palgrave Macmillan and their staff for believing in and supporting this project.

Finally, a personal thank you to family and friends who have supported my endeavors, including this project. Sara, my wife, has given me endless encouragement, support, and love. Thank you for believing in me, your patience, and being our family's rock. Thank you to my children—William, Hope, and Abe—for your love, affection, and joy; to my parents and sister for continual

CONTENTS

Notes on Contributors

Hangsan Paul Ahn is working on his Master's degree in the Department of Communication Arts at the University of Wisconsin-Madison. He is interested in innovation in teams and social science methodology. Before graduate school, he worked at Hyundai Motors in marketing.

Arnoud Arntz, Ph.D. is Professor of Clinical Psychology at the University of Amsterdam. His main research interest lies in the field of personality disorders and includes personality pathology in the forensic field. His research addresses both psychological processes in and treatment of personality disorders.

Kimberly Beach (B.A., Marist College) lives in Connecticut and has worked in the behavioral health field for the last 12 years, connecting individuals to recovery support and clinical services through state and federally funded programs. At present, she facilitates communication between state agencies and community providers to best serve parents and caregivers who are in recovery. Her past research focused on communication aspects of interrogation and torture techniques, which she presented at the National Communication Association's annual conference.

Christian Bender is an undergraduate student studying the Russian language and communication at The University of Alabama. He is currently conducting research on the influences of culture on uses of deception in everyday communication among international students as a Research Assistant in the Human Communication Research Laboratory. Bender intends to study international security in Russia in addition to his current research. He is an avid supporter of the arts and has been cast in several theatrical productions at UA.

Gershon Ben-Shakhar, Ph.D. is Professor Emeritus of Psychology at the Hebrew University of Jerusalem. His main area of research is cognitive psychophysiology focusing on the role of stimulus novelty and significance in

orienting response elicitation and its habituation process, and on psychophysiological detection of deception. He has published more than 120 articles on psychophysiology, as well as on psychological testing, individual differences in cognitive ability, and human judgment and decision making.

Thomas J. Billard is a doctoral student in the Annenberg School for Communication and Journalism at the University of Southern California. He conducts research on transgender coverage in political media, transgender representations in mass media, graphic design in brand culture, and processes of production in graphic design.

Gary D. Bond, Ph.D. is Professor of Psychology at Eastern New Mexico University in Portales, NM.

James Alex Bonus (Ph.D., University of Wisconsin-Madison) is an Assistant Professor in the School of Communication at The Ohio State University. He studies media experiences in childhood. His recent work with young children examines their comprehension of science television, while his work with adults examines the impact of exposure to nostalgic media on well-being.

Melanie Booth-Butterfield (Ph.D., University of Missouri-Columbia) is Professor Emerita in the Department of Communication Studies at West Virginia University. She has emphasized research in the areas of processes and outcomes involved with humor enactments (e.g., humor orientation) and a variety of interpersonal/relational communication issues. She is former editor of *Communication Education* and past President of the Eastern Communication Association.

Joshua Braverman received his B.S. degree in psychology with distinction from Nova Southeastern University in May 2018. He started his Ph.D. in social psychology at the University of Minnesota in Fall 2018.

Gayle Brewer, Ph.D. is a Lecturer at the University of Liverpool. Her research interests focus on the psychology of romantic and sexual relationships. She has published over 60 peer-reviewed journal articles and numerous books and book chapters.

Ray Bull, D.Sc. is Professor of Criminal Investigation at the University of Derby, UK, and immediate past President of the European Association of Psychology and Law.

Zachary M. Carr is currently a doctoral candidate in the Department of Communication at the University at Buffalo, SUNY, under the advisement of Dr. Mark G. Frank. Mr. Carr's research foci capture the intricacies of interpersonal communication and, specifically, nonverbal communication. Additionally, Mr. Carr focuses on the analysis of facial expressions and their role in meaningful dyadic exchanges and deceptive contexts.

Daniella K. Cash is a graduate student working toward her doctorate degree at Louisiana State University. She completed her M.A. degree at the

University of Alabama in Huntsville and her B.S. at Florida Southern College. Daniella is interested in issues pertaining to deception detection, eyewitness accuracy, jury decision making, and other psycho-legal issues.

Innocent Chiluwa, Ph.D. is Professor of Language (Discourse Analysis) and Media/Digital Communications in the Department of Languages & General Studies at Covenant University, Ota, Nigeria. He is a Research Fellow and Digital Expert of the Alexander von Humboldt Foundation (Bonn), and a Visiting Professor in the Department of English, University of Freiburg, Germany. His research interests are discourse analysis and conflict studies; social movement studies; social media and society; cyber civic engagement, online activism, terrorism, and political violence; and online deceptive communications, particularly those associated with phishing, email scams, and cyber-attacks.

Viorela Dan (Ph.D., Free University of Berlin, 2016) is a Postdoctoral Researcher at the Ludwig Maximilian University of Munich. Her research interests include framing theory, the interplay between words and visuals, health communication, and political communication. Her newest publication is *Integrative Framing Analysis: Framing Health Through Words and Visuals* (2018), a monograph published by Routledge.

Coral Dando, Ph.D. is former Police Officer and is Professor of Forensic Psychology at the University of Westminster, UK.

Gary C. David, Ph.D. is Associate Professor of Sociology and Associate Professor of Information Design and Corporate Communication at Bentley University. He conducts ethnographic research in a variety of workplace settings. His past studies include the analysis of interactions between workers and customers in Arab-owned convenience stores in Metropolitan Detroit. He is involved in applying conversation analysis and forensic linguistics to the examination of police interviews, interrogations, and interactions. Other projects include examining the nature of collaborative activity in multicultural worksites, the impact of speech recognition technology and electronic medical records on health care, the implementation of enterprise systems on workplaces, and how co-workers build a collaborative relationship through engaging in workplace practices.

Daniel Cochece Davis (Ph.D., University of Southern California) joined Illinois State University's School of Communication as an Assistant Professor in 2012. His research and teaching examines the neuro-cultural dialectic across all types and contexts of human communication, including torture, polygraphing, and the neurophysiology of deception, as well as areas considered less dark (e.g., intercultural, leadership, and relationships).

Deborah Davis, Ph.D. is Professor of Psychology at the University of Nevada, Reno. She has published a number of articles and chapters on interrogation and confession and has testified or consulted in more than 100 cases involving claims of false confession.

Vincent Denault, LL.M. is a former litigation lawyer and currently a lecturer and Ph.D. candidate in the Department of Communication at the Université de Montréal. His research focuses primarily on issues related to credibility assessment, deception detection, and nonverbal communication during trials. He is also the co-director of Montreal's Center for Studies in Nonverbal Communication Sciences and the inaugural recipient of the Emerging Scholar Award (2016) of the Nonverbal Division of the National Communication Association.

Rachel E. Dianiska is a graduate student working toward her doctorate in cognitive psychology at Iowa State University. She completed her M.S. degree at Iowa State University and graduated with honors from Louisiana State University with a B.S. in psychology and a B.A. in sociology. She is interested in the influence of cognitive processes to issues such as interviewing and credibility assessment and the development of false memories.

Tony Docan-Morgan (Ph.D., University of Washington) is Professor of Communication Studies at the University of Wisconsin-La Crosse. His teaching and research span the topics of public communication, interpersonal relationships, and ethics and deception. As a proponent of applied scholarship and civic engagement, his work aims to help others become effective, empowering, and ethical communicators. He has helped North Korean refugees tell their stories of survival, escape, and readjustment.

James E. Driskell, Ph.D. is the President and Senior Scientist at Florida Maxima Corporation, an Adjunct Professor of Psychology at Rollins College, and Fellow of the American Psychological Association. Dr. Driskell has served as principal investigator on research projects related to deception for the Department of Homeland Security, Defense Intelligence Agency, Federal Bureau of Investigation, and other organizations.

Tripp Driskell, Ph.D. is a Research Scientist with Florida Maxima Corporation in Orlando, Florida. He received a Ph.D. in applied experimental and human factors psychology from the University of Central Florida and has conducted recent research on deception and investigative interviewing for the Federal Bureau of Investigation.

Norah E. Dunbar, Ph.D. is a Professor of Communication at the University of California Santa Barbara. She teaches courses in nonverbal and interpersonal communication, communication theory, and deception detection. She is also an affiliate faculty member in the Center for Information Technology & Society and the Center for Digital Games Research. She has received over $13 million in external research funding from US federal agencies, published over 60 peer-reviewed journal articles and book chapters, and presented over 100 papers at national and international conferences. She is on the editorial board of six disciplinary journals and served as the Chair of the Nonverbal Division of the National Communication Association from 2014–2016.

Eitan Elaad, Ph.D. is Professor of Psychology in the Department of Behavioral Sciences at Ariel University, Israel. Between 2006 and 2009, he was the chairman of the department. Since 1977, he has been engaged in research and lectures on lying and lie detection. His main fields of interest are detection of deception with special emphasis on psychophysiological detection of concealed information, the psychology of lying, decision making and biases, and legal psychology.

Elaine E. Englehardt, Ph.D. is the Distinguished Professor of Ethics and Professor of Philosophy at Utah Valley University. She has written or edited over ten books on ethics and written numerous journal and encyclopedia articles. She has been the principal investigator of a variety of national grants in ethics and was a senior administrator at Utah Valley University for more than 12 years.

Maria Espinola, Psy.D. is Assistant Professor at the Department of Psychiatry and Behavioral Neuroscience at the University of Cincinnati College of Medicine. She is a licensed clinical psychologist, and her areas of clinical and research focus are multicultural psychology, women's issues, and trauma psychology. She completed her doctorate in clinical psychology at Nova Southeastern University, her predoctoral fellowship in multicultural psychology at Boston University Medical Center, and her postdoctoral fellowship in trauma psychology at McLean Hospital and Harvard Medical School.

Louis Fisher, Ph.D. is Scholar in Residence at the Constitution Project and Visiting Scholar at the William and Mary Law School. For four decades, he served at the Library of Congress as Senior Specialist in Separation of Powers (1970–2005) and Specialist in Constitutional Law at the Law Library (2005–2010). Author of 26 books and over 500 articles, he has testified more than 50 times before congressional committees on a range of constitutional issues. His most recent book is *President Obama: Constitutional Aspirations and Executive Actions* (2018).

Annette L. Folwell (Ph.D., University of Oklahoma) is an Associate Professor in the Department of Psychology and Communication Studies at the University of Idaho. Her research centers on interpersonal communication within the family, particularly marital, sibling, and grandparent–grandchild relationships, as well as instructional communication. Dr. Folwell's research has been published in *Journal of Social and Personal Relationships, Journal of Language and Social Psychology*, and *Journal of the Northwest Communication Association*.

Mark G. Frank, Ph.D. is Professor and Chair in the Department of Communication at the University at Buffalo, SUNY. Dr. Frank received his Ph.D. from Cornell University and was a postdoctoral research fellow at the University of California San Francisco School of Medicine, under the tutelage of Dr. Paul Ekman. Dr. Frank's primary research interests include the study of

nonverbal communication, with a focus on understanding the complexities of facial expressions and deception in meaningful, real-world settings.

Mar Gámez García, M.A. is currently a Ph.D. student and teaching assistant in the Department of Romance Languages and Literatures at the University of Cincinnati. She holds an M.A. in Spanish from the University of Cincinnati, a B.A. in Journalism from University Carlos III de Madrid, and a B.A. in English Philology from the University of Barcelona. She wrote the play *La Fauna del Poder* (2016) and has published several peer-reviewed articles on Spanish and Latin American literature.

Kylie L. Geiman (M.S., Purdue University, 2017) is employed as a data analyst in private business.

Mark A. Generous (Ph.D., Arizona State University, 2016) is an Assistant Professor of Communication at California State University Polytechnic, Pomona. His research focuses on close, interpersonal relationships; specifically, he examines difficult and challenging conversations within the context of intimate relationships. His work has focused on end-of-life communication, sexual communication, and profanity.

Linda Marjoleine Geven is a joint doctoral candidate at the University of Amsterdam, The Netherlands, and the Hebrew University of Jerusalem, Israel. As a legal psychologist, her current research on memory detection focuses specifically on bridging the gap between research and practice. Moreover, she is passionate about investigating and preventing miscarriages of justice.

Pär Anders Granhag, Ph.D. is a Professor of Psychology at the University of Gothenburg, Sweden. He has been conducting research on legal psychology for almost 30 years. His main areas of research are interviewing, interrogation, and deception detection. He is the Founding Director of the Research Unit for Criminal, Legal and Investigative Psychology (CLIP), and he is the Editor of the journal *Applied Cognitive Psychology*.

John O. Greene (Ph.D., University of Wisconsin-Madison, 1983) is a Professor in the Brian Lamb School of Communication at Purdue University. He is a former editor of *Human Communication Research* (2001–2003) and former book review editor of *Communication Theory* (1993–1996). He is a recipient of the National Communication Association's Charles H. Woolbert Research Award (1994) and a two-time recipient of the Gerald R. Miller Book Award (2002, 2004).

Darrin J. Griffin (Ph.D., Buffalo SUNY, 2014) is an Assistant Professor at The University of Alabama in Tuscaloosa where he serves as the Director of the Human Communication Research Laboratory. Griffin conducts research on communication topics related to nonverbal communication, deception, and deafness. His teaching schedule most often includes an undergraduate

course on truth, ethics, and deception, and he has taught this course as a doctoral seminar as well. Griffin is a hearing person, a child of deaf adults (CODA), and identifies as a member of the deaf community; he is fluent in American Sign Language. His bilingual-bicultural status allows him to use his teaching and research to be an advocate for deaf people. More information about Griffin can be found on his website www.darringriffin.com.

Jeffrey T. Hancock (Ph.D., Dalhousie University) is a Professor in the Department of Communication at Stanford University. He works on understanding psychological and interpersonal processes in social media by using computational linguistics and behavioral experiments to examine deception and trust, emotional dynamics, intimacy and relationships, and social support.

Sean M. Horan (Ph.D., West Virginia University) is Associate Professor and Chair of the Department of Communication at Fairfield University. His research primarily centers upon Affection Exchange and Work-Life Border theories, specifically through investigations of sex, deceptive affection, and workplace romance. He writes regularly for *Psychology Today* and is an associate editor of the *Journal of Social and Personal Relationships.*

Anne P. Hubbell (Ph.D., Michigan State University, 2000) is Professor of Communication at New Mexico State University in Las Cruces, NM. Her teaching and research are in the areas of health and organizational communication. Within organizational communication, Dr. Hubbell focuses on deception with a particular emphasis on the role of trust and justice in reducing deceptive behaviors.

Neslihan James-Kangal, M.A. is a Clinical Psychology doctoral student at the University of Cincinnati. For her Master's thesis, Neslihan evaluated couples' accuracy in detecting deception and trained them in a system of investigative interviewing to improve their accuracy. She plans to pursue a career in academia and continue her research on how couples, particularly young dating couples, handle conflict in their relationships.

Pamela J. Kalbfleisch (Ph.D. Michigan State University, 1985) is Professor of Communication at the University of North Dakota. She is also a member of the General and Experimental Psychology Graduate Faculty at the University. Dr. Kalbfleisch's research focuses on deception, issues of trust and distrust in relationships, and mentoring relationships. Kalbfleisch is the author of *Mentoring Enactment Theory.* She is also personally dedicated to helping members of underrepresented groups find mentors and sources of social support.

Trevor Kauer (B.S., University of Idaho) is a graduate student and teaching assistant in Communication Studies at Texas State University. His prior experience as an Emergency Medical Technician informs his research interests on the ways our understandings about interpersonal, health, and family communication apply to first responders and public safety employees.

Jode Keehr (B.S., University of Idaho) is a doctoral student in the experimental Psychology Program at the University of Idaho. She is a professional consultant on electronic media presence for businesses and, with that expertise, developed a course for the University of Idaho on social messaging for small organizations. Her research interests in human systems integration reside at the intersection of technology, applied cognition, and cognitive neuroscience.

Megan Kenny is a Lecturer in Forensic Psychology at Sheffield Hallam University. Her research interests include all areas of interpersonal crime, particularly domestic and sexual violence as well as the use of technology to facilitate offending behavior. She has worked on the multinational None in Three project and has conducted research exploring sexually motivated abduction. Her Ph.D. thesis is entitled *Who's Following You? Cyberviolence on Social Media*.

Merel Kindt, Ph.D. is Professor of Experimental Clinical Psychology in the Department of Psychology at the University of Amsterdam. Over the years, she has established a highly successful program of research on the plasticity of fear memory. Specifically, she has convincingly demonstrated that the emotional expression of fear memories can be neutralized in humans and which conditions are essential for this to occur. The potential of her research program lies in the unique bidirectional translational approach: It builds on fundamental insights from animal and human neuroscience literature as well as on clinical science.

Bennett Kleinberg is a doctoral researcher at the University of Amsterdam. His research is on large-scale airport security innovation at the intersection of cognitive psychology, crime science, computational linguistics, and artificial intelligence.

Mark L. Knapp (Ph.D., Pennsylvania State University) is the Jesse H. Jones Centennial Professor Emeritus in the Department of Communication Studies and Distinguished Teaching Professor Emeritus at The University of Texas at Austin. He has published several works on lying and deception. He co-edited the *Handbook of Interpersonal Communication* and *The Interplay of Truth and Deception*, and co-authored *Lying and Deception in Human Interaction*, *Nonverbal Communication in Human Interaction*, and *Interpersonal Communication in Human Relationships*. He is a past President and Distinguished Scholar of the National Communication Association and a past President and Fellow of the International Communication Association.

Sean M. Lane, Ph.D. is Professor of Cognitive and Brain Sciences in the Department of Psychology at Louisiana State University. His research examines both basic and applied aspects of learning and memory and has been funded by the National Science Foundation, the Federal Bureau of Investigation, the Spencer Foundation, and the Louisiana Board of Regents. Besides his work in lying and memory, his other research interests include

eyewitness suggestibility, memory beliefs, the impact of emotional arousal on memory, and the application of findings from learning and memory research to educational settings.

Samuel Larner, Ph.D. is a Senior Lecturer in Forensic Linguistics in the Department of Languages, Information and Communications at Manchester Metropolitan University, UK. His research interests lie primarily in the use of linguistics for investigative and evidential purposes, including deception detection, the theory and practice of forensic authorship analysis, and the language children and young people use when disclosing sexual abuse.

Seow Ting Lee (Ph.D., University of Missouri-Columbia) is Professor in the University of Colorado Boulder's Department of Advertising, Public Relations and Media Design. Her research in strategic communication focuses on organization–stakeholder relationships and communication strategies with an emphasis on ethics and new media. Her research has been published in the *Journal of Communication*, *Public Relations Review*, *Journal of Public Relations Research*, *Journal of Health Communication*, *Health Communication*, *Media Culture and Society*, *Mass Communication and Society*, *Health Promotion International*, and *Journal of Mass Media Ethics*.

Paul Martin Lester, Ph.D. is a Clinical Professor for the School of Arts, Technology, and Emerging Communication at the University of Texas at Dallas and a Professor Emeritus in Mass Communications at California State University, Fullerton. He received a bachelor of journalism from the University of Texas at Austin, a master's from the University of Minnesota, and a Ph.D. from Indiana University. He is the author or editor of several books that include: *Visual Ethics: A Guide for Photographers, Journalists, and Filmmakers* (2018), *Visual Communication Images with Messages Seventh Edition* (2017), *Digital Innovations for Mass Communications Engaging the User* (2014), *Images that Injure Pictorial Stereotypes in the Media Third Edition* with Susan Ross (2011), *Visual Journalism: A Guide for New Media Professionals* with Christopher Harris (2002), *Desktop Computing Workbook: A Guide for Using 15 Programs in Macintosh and Windows Formats* (1996), and *Photojournalism: An Ethical Approach* (1991). From 2006 until 2011, he was editor of *Visual Communication Quarterly*, and from 2011 until 2015, he was editor of the AEJMC publication *Journalism and Communication Monographs*.

Timothy R. Levine, Ph.D. is Distinguished Professor and Chair of Communication Studies at University of Alabama at Birmingham. Levine teaches and conducts research on topics related to interpersonal communication, persuasion and social influence, cross-cultural communication, and social scientific research methods. Levine has published more than 125 refereed journal articles reporting original research related to communication, and he is an internationally recognized leader in deception research. He is the author of *Truth Default Theory*.

Yachao Li (M.A., Rutgers University, 2015) is a doctoral candidate in the Department of Communication Studies at the University of Georgia. His research focuses on interpersonal and health communication in the contexts of same-sex relationships, specifically how same-sex couples and their families manage relational challenges. He also teaches introduction to interpersonal communication, advanced interpersonal communication in health contexts, and small group communication.

Tyler N. Livingston is a Ph.D. student in the Interdisciplinary Social Psychology Program at the University of Nevada, Reno. He is specifically interested in the study of social cognition and human memory.

Eric Mac Giolla, Ph.D. is a postdoc at the University of Gothenburg, Sweden. He has a broad interest in legal psychology and has conducted research on the topics of eyewitness memory, investigative interviewing, and deception detection.

David M. Markowitz, Ph.D. is an Assistant Professor in the School of Journalism and Communication at the University of Oregon. He uses language data from natural repositories to make inferences about people; what they are thinking, feeling, and experiencing psychologically. A large part of his research focuses on how deception affects language, such as how fraudulent scientists write their research papers compared to genuine scientists. He received his Ph.D. from Stanford University and his undergraduate and master's degrees from Cornell University.

Jaume Masip, Ph.D. is an Associate Professor in the Department of Social Psychology and Anthropology at the University of Salamanca, Spain, where he currently teaches legal psychology and criminal psychology. He has conducted extensive research on nonverbal and verbal detection of deception, with a particular (but not exclusive) focus on police officers' detection abilities, response biases, and beliefs about deception cues. He has given talks and workshops on deception detection to law enforcement officers and other practitioners.

Matthew S. McGlone (Ph.D., Princeton University) is Professor of Communication Studies at The University of Texas at Austin. His research investigates persuasion, deception, and stereotyping in social interaction. His research efforts have been funded by the National Science Foundation, US Departments of Education and Justice, and the Kozmetsky Family Foundation. With Mark L. Knapp, he co-edited *The Interplay of Truth and Deception* (2010, Routledge) and co-authored *Lying and Deception in Human Interaction* (2016, Kendall Hunt). He is Associate Director of UT's Center for Health Communication.

Christian A. Meissner, Ph.D. is Professor of Psychology at Iowa State University. He conducts empirical studies in applied cognition, including the role of memory, attention, perception, and decision processes in real-world

tasks. He has published more than 80 peer-reviewed journal articles and book chapters and has received more than $15 million in grant funding from such agencies as the National Science Foundation, the US Departments of Defense, Justice, and Homeland Security, and the US Intelligence Community.

Taylor Meredith graduated Summa Cum Laude from Lake Superior State University in May 2017 with concentrations in psychology and criminal justice. She has held several positions in human services and plans to attend graduate school in the forensic behavioral sciences.

David Miller is Professor of Political Sociology in the School of Policy Sciences at the University of Bristol. From 2013 to 2016, he was RCUK Global Uncertainties Leadership Fellow leading a project on understanding and explaining terrorism expertise in practice. David is a co-director of Public Interest Investigations, a nonprofit company of which Spinwatch. org and Powerbase.info are projects and co-director of the Organisation for Propaganda Studies. Recent publications include "Organized Persuasive Communication: A New Conceptual Framework for Research on Promotional Culture, Public Relations and Propaganda" in *Critical Sociology* (co-author, 2018); *Impact of Market Forces on Addictive Substances and Behaviours: The Web of Influence of Addictive Industries* (co-author, 2018, Oxford University Press); and *What Is Islamophobia? Racism, Social Movements and the State* (co-editor, 2017, Pluto Press).

Wendy L. Morris (Ph.D., University of Virginia, 2005) is Professor of Psychology and Dean of the Faculty at McDaniel College. She is a social psychologist, and her two areas of research are deception within friendships and stereotypes, prejudice, and discrimination.

Marley Morrow received her M.S. degree in Experimental Psychology from Nova Southeastern University in June 2016. Currently, she is a Senior Marketing Manager and Digital Content Editor at Impact Analytics, Inc.

Kate Muir, Ph.D. is a Research Associate in the School of Management at the University of Bath. Kate is a Chartered Psychologist and Chartered Scientist with the British Psychological Society. Her research interests include face-to-face and computer-mediated forms of human communication, and social and personality influences upon communication behaviors and autobiographical memory.

Galit Nahari, Ph.D. is a Professor and the head of the M.A. research track in the Department of Criminology at Bar-Ilan University. She is the Editor of *Legal and Criminological Psychology*. Her main research interest is legal and investigative decision making. As part of it, she develops interrogation and lie-detection methods and examines their validity and suitability in different contexts, and their vulnerability to judgmental biases. She lectures and

conducts training workshops on verbal lie detection for academics and practitioners in Europe and Israel.

Natalie Odisho, B.S. is an Oriental Medicine master's student at Florida College of Integrative Medicine. She earned her bachelor's degree in media communication studies with a minor in psychology at the Florida State University in 2013. As an artist and Bright Futures scholar, Natalie Odisho investigates the relationships between art and health.

Lening Alexis Olivera-Figueroa, Psy.D., ABPP is a board-certified clinical psychologist licensed in Connecticut, Puerto Rico, and Quebec, Canada. Dr. Olivera-Figueroa completed his doctorate in clinical psychology at Ponce Health Sciences University in Ponce, Puerto Rico, his predoctoral fellowship at the National Institute of Mental Health, and his five postdoctoral fellowships at McGill University, University of Montreal, Yale University, the Veteran Affairs-Connecticut Healthcare System, and the Syracuse Veterans Affairs Medical Center. In addition, Dr. Olivera-Figueroa has held faculty appointments at Yale University, Rutgers University, and the Ana G. Méndez University System-Capital Area Campus in Wheaton, Maryland.

Evan Polman (Ph.D., Cornell University) is Assistant Professor of Marketing at the Wisconsin School of Business, University of Wisconsin-Madison. He carried out his postdoctoral work at New York University. His research focuses on decision making, especially how people make decisions and judgments that involve others.

John H. Powers (Ph.D., University of Denver, 1977) is Professor Emeritus of Communication Studies at Hong Kong Baptist University, where he has been for the nearly 25 years. Previously, he taught in the Speech Communication Department at Texas A&M University in the USA. His research interests include paradigm building in the communication discipline, theory and criticism of public discourse, and the role of language in routine communication practices. He is editor, with Randy Kluver, of *Civic Discourse, Civil Society, and Chinese Communities* (Ablex, 1999), author of *Public Speaking: The Lively Art* (HarperCollins, 1994; McGraw-Hill, 2016), and editor, with Xiaosui Xiao, of *The Social Construction of SARS: Studies of a Health Communication Crisis* (Amsterdam and Philadelphia: John Benjamins, 2008). He is currently developing three interconnected projects: *Principles of Human Communication: A Tier-Based Introduction*, *Rhetorical Criticism of Public Communication: Discursive and Presentational Dimensions*, and *Toward a New Paradigm for the Human Sciences: A User's Guide to Susanne Langer's Philosophy of Mind*.

Danielle Blake Prentice (M.A., San Diego State University) currently lives in San Diego and previously completed the Marine Corps' Officer Candidate School. She has researched communication aspects of interrogation and torture techniques, as well as health communication, and has presented her work

at the National Communication Association's annual conference. She has also consulted with school districts and colleges to recover federal reimbursement dollars for campus medical services. In 2013, she decided to become a stay-at-home mom.

Michael S. Pritchard, Ph.D. is Emeritus Professor of Philosophy and former Director of the Center for the Study of Ethics at Western Michigan University. He is author or co-author of several books on ethics, including *Communication Ethics*, 2nd ed. (Wadsworth, 1994), co-authored with James A. Jaksa. He and Elaine E. Englehardt also served as co-editors of the journal *Teaching Ethics*.

Douglas E. Pruim (Ph.D., Purdue University, 2018) is a Clinical Associate Professor of Managerial Communications in the Krannert School of Management at Purdue University. His research focuses on adult communicative play.

Richard Reardon, Ph.D. is a Professor of Psychology and Organizational Sciences at the University of Idaho. He received his training from the University of North Carolina at Chapel Hill (A.B.) and the University of Georgia (M.S. and Ph.D.) in social psychology. He has published in the areas of memory source monitoring, cognitive generation, deception, and marketing communication. Reardon has also held various university administrative appointments at Idaho and the University of Oklahoma, where he spent the first half of his career.

Peter O. Rerick is a Ph.D. student in the Interdisciplinary Social Psychology Program at the University of Nevada, Reno. He is specifically interested in the study of social psychology and law.

Piers Robinson is Chair in Politics, Society and Political Journalism, School of Social Sciences, at the University of Sheffield. He researches international politics, communications, and contemporary propaganda with a particular focus on war and conflict. He is co-director of the *Organisation for Propaganda Studies* and convenes the *Working Group on Syria, Media and Propaganda*. He is author of *The CNN Effect: The Myth of News Foreign Policy and Intervention* (Routledge, 2002), co-author of *Pockets of Resistance: British News Media, War and Theory in the 2003 Invasion of Iraq* (Manchester University Press, 2010) and co-editor of the *Routledge Handbook of Media, Conflict and Security* (2017).

Nicholas O. Rule, Ph.D. is an Associate Professor of Psychology and Canada Research Chair in Social Perception and Cognition at the University of Toronto, in Toronto, Canada.

H. Russell Searight, Ph.D. is Professor of Psychology at Lake Superior State University. He received a Ph.D. in Clinical Psychology from Saint Louis University and an MPH from Saint Louis University School of Public

Health. Prior to his current position, Searight was Clinical Associate Professor of Community and Family Medicine at Saint Louis University School of Medicine and Director of Behavioral Medicine at the affiliated Forest Park Hospital Family Medicine Residency program. During that time, he also served as both member and chair of the hospital's medical ethics committee. His recent areas of scholarly interest include mental health practice in primary health care, cross-cultural issues in mental health and medicine, and medical ethics.

Jennifer A. Samp (Ph.D., University of Wisconsin-Madison, 1999) is a Professor of Communication at University of Georgia. Dr. Samp's work illuminates how and why individuals do not always respond the same way when managing relational problems and conflicts with close friends, romantic partners, and family members. Dr. Samp is a Fellow of the UGA Owens Institute for Behavioral Research, a Faculty Affiliate of the UGA Center for Risk Communication, and a Faculty Affiliate of the Emory University Center for Injury Control. Her research has been supported by grants from the National Institutes of Health, UGA Research Foundation, UGA Owens Institute for Behavioral Research, the Arthur W. Page Center, and the National Endowment for the Humanities. She is currently Chair of the Interpersonal Communication Division of the International Communication Association (ICA) and editor of the *Southern Communication Journal.*

Sara M. Schewe received a B.S. in Communication Disorders from Eastern New Mexico University in Portales, NM, and is a master's student in communication sciences and disorders at University of Wisconsin-Eau Claire.

R. Scott Tindale, Ph.D. is a Professor of Social Psychology at Loyola University of Chicago. His main research areas are individual and group decision making and group performance. His most recent work focuses on how groups and individuals differ when facing decisions with ethical implications. He also serves as an associate editor for *Group Processes and Ingroup Relations.*

Kim B. Serota, Ph.D. is a Special Instructor in Marketing at Oakland University. As a veteran of more than 30 years in automotive marketing management, Serota teaches marketing research, marketing theory, integrated marketing communications, and management strategy. With a background in communication theory, his research interests include the prevalence of lying, the application of complexity theory to lie behavior, deception detection, and marketing deception.

David Shulman, Ph.D. is a Professor of Anthropology and Sociology at Lafayette College. He has published five books and various articles addressing aspects of deception, impression management, marketing, and social psychology.

Angelica Snyder received a B.S. in Psychology from Eastern New Mexico University in Portales, NM, and is a master's student in mental health counseling at University of the Southwest.

Anne Solbu, Ph.D. is a Postdoctoral Research Associate in the Department of Communication at SUNY at Buffalo, where she also received her doctoral degree. Her area of study encompasses human nonverbal behavior with an emphasis on contextual assessment. She is involved in research on deception, including behavior while lying, deceptive intent, beliefs about lying, and deception detection. She is also interested in self-awareness and its implications on expression of emotion and behavior in general.

Sarah A. Son, Ph.D. is Research Director at the Transitional Justice Working Group, a Seoul-based NGO focused on documenting human rights abuses in North Korea. She is also a Research Professor at Korea University's Human Rights Center. Son was awarded her Ph.D. by the School of Oriental and African Studies at the University of London in 2014. Her research interests include North-South Korean relations, North Korean refugees, migration, identity, nationalism, transitional justice, and post-conflict peacebuilding.

Lassiter F. Speller received a M.S. in Psychology from The Ohio State University in Columbus, OH, and is a Ph.D. student in Cognitive Psychology at The Ohio State University.

Susan A. Stearns, Ph.D. is a Professor of Communication Studies at Eastern Washington University. She focuses her research interests in the scholarship of learning, Erving Goffman, deception, academic integrity, and organizational communication. A common theme through these diverse areas is discovering information leading to problem resolution.

R. Weylin Sternglanz (Ph.D., University of Virginia, 2003) is an Associate Professor of Psychology at Nova Southeastern University. He is a social psychologist whose principal research focus is in the field of interpersonal perception. Primarily, he studies deception detection and the ability to decode nonverbal cues to emotions.

Chris N. H. Street, Ph.D. is a Senior Lecturer at the University of Huddersfield, UK, and Director of the Social Conflict Lab. His work takes a cognitive theoretical perspective with an aim to understand how lie–truth judgments are formed by making use of behavioral experiments, eye tracking, and computational modeling. He is author of the Adaptive Lie Detector theory (ALIED) and is currently working to develop a computational model of ALIED under the ACT-R cognitive architecture.

Tara Suwinyattichaiporn (Ph.D., Arizona State University, 2016) is an Assistant Professor at California State University, Fullerton. Her research emphasis is on the intersection of intercultural communication and personal/

social relationships. Her work has been presented at the National and International Communication Association conferences and appeared in *Communication Teacher* and *Basic Course Annual*.

Golnar B. Teimouri is a graduate of the M.A. Program in Social Sciences at the University of Chicago. For her M.A. thesis, she examined how members of congress portray U.S. mass shootings on their Twitter accounts. She also earned a B.A. from the University of California, Berkeley with a major in sociology and minor in human rights. Her work in the field of political science continues to focus on U.S. American politics. She currently works on a political campaign in Chicago.

Catalina L. Toma (Ph.D., Cornell University) is an Associate Professor in the Department of Communication Arts at the University of Wisconsin-Madison. Her research examines how people understand and relate to one another when interacting via communication technologies (e.g., online dating, social network sites, and texting). She focuses on relational processes such as self-presentation, deception, and emotional well-being.

James Trainum is a retired detective from the Metropolitan Police Department of Washington, D.C., where he spent the majority of his career in the homicide branch. He currently consults for Criminal Case Review and Consulting on questions regarding police practices, specializing in interview and interrogation tactics and the evaluation of statement evidence. Other projects include developing training for lawyers, judges, and police officers on identifying and avoiding pitfalls common to investigations. His research also looks into developing root cause analysis processes by which the criminal justice system can learn from investigative failures.

Maureen van der Burgh is a Researcher and Forensic Psychologist at the Diagnostic Expertise Centre in the Netherlands.

Lyn M. Van Swol (Ph.D., University of Illinois at Urbana-Champaign) is a Professor in the Department of Communication Arts at the University of Wisconsin-Madison. Her research examines information sharing in groups, advice utilization, and deception.

Bruno Verschuere, Ph.D. is Associate Professor of Forensic Psychology at the University of Amsterdam. He regularly gives workshops and lectures on lie detection, has served as expert witness on lie detection in court, and has published over 50 papers on deception and lie detection. He is the co-founder of the European consortium of Psychological Research On Deception Detection (E-PRODD; http://eprodd.eu/). His latest book, *Detecting Deception: Current Challenges and Cognitive Approaches* edited together with Aldert Vrij and Pär Anders Granhag, appeared with Wiley in 2015.

J. Guillermo Villalobos is a Ph.D. candidate at the Interdisciplinary Social Psychology Program at the University of Nevada, Reno. He is interested in the psychological study of vulnerable populations in the criminal justice system, as well as in policy advocacy and reform.

Cynthia Adarian Wartanian (B.A., Marist College) spent four years traveling the USA and abroad representing a luxury jewelry brand before shifting to the nonprofit community and managing fundraising events for underserved health conditions like blood cancers and epilepsy. In 2013, she decided to return to Long Island and the luxury jewelry industry where she focuses in ecommerce merchandising. Cynthia's past research focuses on the communication aspects of interrogation and torture techniques, and has presented this work at the National Communication Association's annual conference.

Emma J. Williams, Ph.D. is a Vice Chancellor's Fellow in Digital Innovation and Wellbeing in the School of Experimental Psychology at the University of Bristol. Emma is a chartered psychologist with the British Psychological Society and a chartered scientist with the Science Council. Her research interests include online deception, susceptibility to online influence, and human aspects of cyber security.

John Paul Wilson, Ph.D. is an Assistant Professor of Psychology at Montclair State University in Montclair, New Jersey.

Jeremy R. Winget is a graduate student of Social Psychology at Loyola University Chicago, where he is currently working toward his Ph.D. His main research interests focus on the factors that enhance or limit cooperation and conflict via group dynamics. His most recent work focuses on comparing individual and group decision making within moral dilemmas.

William Douglas Woody, Ph.D. is Professor of Psychological Sciences at the University of Northern Colorado, where he teaches and conducts research on psychology and law, history of psychology, and teaching of psychology. His scholarship includes both experimental and historical studies of interrogation and confession.

Marjorie Yambor, Ph.D. is an Associate Professor for the Department of Communication at Dalton State College in Georgia. She received a B.A. from Western Kentucky University, an M.A. from University of Alabama, and a Ph.D. from Michigan State University. In addition to her published work with Dada and violent artistic expression, she presents annual panels at the Broadcast Education Association and serves on the editorial board for *Visual Communication Quarterly*. Her experience in the advertising industry, college radio, and higher education—environments rich in their unique nuances of deceptive communication—has guided her research interests in popular culture, visual communication, cultivation, and gender.

LIST OF FIGURES

LIST OF TABLES

An Introduction and Overview of Deceptive Communication

Historical Perspectives on the Study of Lying and Deception

Matthew S. McGlone and Mark L. Knapp

When philosopher George Santayana (1906) said "history is always written wrong and so needs to be rewritten" (p. 397), he was reminding us that events can be seen differently by different people and the way one views them may vary at different points in time. It doesn't mean every interpretation is equally meritorious; just that there can be different versions. In this chapter, we are writing the history we know coupled with the knowledge that in today's globally interconnected world there is always more to know. The irony that historical truth may have some wiggle room when the subject is lying and deception has not gone unnoticed. Keeping this in mind, we humbly forge ahead—identifying what we believe to be key developments or milestones that have shaped the study of lying and deception.

The history of our very human tendency to use verbal and nonverbal behavior to mislead others is probably as old as the species itself. So throughout human history, the subject of deceit has been of special interest to those concerned with establishing the rules of conduct for social behavior. Some of these accounts that span history will be identified as a preface to our primary focus—the rapid growth of academic studies focused on the interpersonal behavior of liars and the detection of lying during the past half century.

M. S. McGlone (✉) · M. L. Knapp
Department of Communication Studies,
The University of Texas at Austin, Austin, TX, USA
e-mail: matthew_mcglone@austin.utexas.edu

M. L. Knapp
e-mail: mlknapp@austin.utexas.edu

ANTIQUITY THROUGH THE MIDDLE AGES

It is in our nature to mislead, but we also generally dislike being misled, with a few exceptions (art, fiction, magic, surprise parties, etc.). Consequently, human societies throughout history have punished those who deliberately mislead and treated the question of how to detect misleaders as serious food for thought. In fact, some of the earliest recorded thoughts about deception detection also happen to be about food. In the sacred Hindu *Yajurveda*, written around 1000 BCE, instructions are offered to royalty for detecting spies disguised as servants with the intent of poisoning food:

> A person who intends to poison food may be recognized. He does not answer questions, or they are evasive answers; he speaks nonsense, rubs the great toe along the ground, and shivers; his face is discolored; he rubs the roots of the hair with his fingers; and he tries by every means to leave the house. (Chand, 1980, p. 54)

In ancient China, food itself was used as a detection tool. People suspected of lying were sometimes forced to chew dry rice while listening to the accusations against them. Afterward, the expectorated rice was examined, and if deemed too dry, it was considered evidence of guilt (Ford, 2006). Greek biographer Plutarch relates an episode in which celebrated anatomist and physician Erasistratus (300–250 BCE) treated food aversion as a deception cue. The physician was asked by Syrian king Seleucus Nicator to examine his son Antiochus, who had stopped taking meals and was wasting away. Noting the prince's mysterious illness started right after his widower father had taken a beautiful new bride, Erasistratus suspected the cause was the prince's efforts to conceal from the king his infatuation with the new queen. Reinforcing this conclusion was his observation that during a physical examination, the prince's pulse quickened when she entered the room (Trovillo, 1939). In each of these ancient cases, the strategy for detecting deception was predicated on an assumption that the somatic expression of anxiety (shifting posture, dry mouth, loss of appetite, heightened pulse, etc.) is evidence of deceptive intent. Despite numerous conceptual and empirical challenges to this assumption, it has continued to dominate deception research and theory in the modern era (Knapp, McGlone, Griffin, & Earnest, 2016).

The subject of deception was integral in the establishment and prescribed conduct of all major religions. Truthfulness is generally recommended in the Quran, but it also says Muslims are permitted to lie, especially to non-believers, if the lie benefits Islam. Judaism and Christianity placed deception at the front and center of their narratives of the human condition. As historian Dallas Denery (2015) aptly put it, the Book of *Genesis* tells us "it took God six days to create the world and the Devil two deceptive sentences to undo it" (p. 21). Christian theology often advocated the belief that no lie was permitted among believers. In his treatise *De Mendacio* ("On Lying"),

St. Augustine of Hippo (354–430) famously argued that lying violates religious precepts and, like a disease, infects personality and destroys integrity (Muldowney, 2002). Living for truth, no matter what the consequences, is to live for God. Augustine not only argued for the moral bankruptcy of lying, but also went on to say that God is perfectly capable of extricating from trouble those who stand fast in the truth. Many subsequent and influential theologians embraced Augustine's absolutist stance, such as St. Thomas Aquinas (1225–1274) and John Wesley (1703–1791), as well as moral philosopher Immanuel Kant (1724–1804).

Augustine's treatise also provided a Christian rationale for the "trial by ordeal," in which an accused liar's innocence was tested by a physical challenge. Importantly, this detection strategy was based not on any theory about the psychophysiological processes associated with stress or guilt, but purely on religious faith. If the accused were innocent, according to this reasoning, God would intervene and protect them from harm. Although trials by ordeal existed long before the Common Era, medieval clergy were particularly creative in devising painful plights from which only divine intervention could spare the innocent (Trovillo, 1939). Trials by fire required the accused to walk over molten coals, touch their tongues to red-hot pokers, or remove a stone from a boiling cauldron; trials by water submerged suspects rope-bound and headfirst into a cold stream and acquitted only the survivors (unless accused of sorcery—witches were thought able to float); trials by poison forced defendants to swallow nightshade and live through the dangerous fever that followed; etc. However, the clergy reserved for accused liars within their ranks the trial by ingestion, or *Corsnaed*, which hardly seems to count as an ordeal (Mackay, 1857):

> A piece of barley bread and a piece of cheese were laid upon the altar, and the accused priest, in his full canonicals and surrounded by all the adjuncts of Roman ceremony, pronounced certain conjurations, and prayed with great fervency for several minutes. The burden of the prayer was that if he were guilty of the crime laid to his charge, God would send the angel Gabriel to stop his throat, and he might not be able to swallow the bread and cheese. There is no instance upon record of a priest having been choked in this manner. (p. 314)

The Renaissance

By the thirteenth century, trials by ordeal for suspected liars were being supplanted by an inquisitorial system of criminal justice across Europe. To obtain a confession from a suspected liar, magistrates were authorized to use torture in their pursuit of judicial certainty. But even in this early, cruel phase of jurisprudence, legal scholars recognized that the practice exposed potentially innocent suspects to unjust suffering and had the potential to produce false confessions (Langbein, 1977). The use of judicial torture eventually declined in the early seventeenth century, when late Renaissance thinkers

advanced a more probabilistic appraisal of the trustworthiness of human testimony with a margin of uncertainty (Andrews, 1994). Increasingly, suspects and witnesses were probed via cross-examination by lawyers, although magistrates still retained ultimate say about the veracity of their testimony. At about the same time, experts began appearing in courts to evaluate various forms of circumstantial evidence (crime scene details, suspect temperament, etc.) that presumably could not be directly assessed by magistrates or jury members, but could be used to corroborate or challenge suspect testimony. These practices have persisted into the present day in Western Europe and the US.

While strategies for unmasking lies evolved in criminal courts during the late Middle Ages and Renaissance, royal courts during this period allowed deceptive strategems to flourish. In the *Politcraticus*, English courtier John of Salisbury (1159/1990) bitterly criticized the royal courts of Europe as infested with "flatterers, wheedlers, gift-givers, actors, mimics, procurers, and gossipmongers" who had rendered the world "treacherous for men of honest virtue."

Ambitious courtiers seeking fortune and power had to maintain a difficult balance, slandering competitors on the one hand while flattering superiors on the other. They typically justified their deceit with time-honored "when in Rome" logic. In an environment where any seemingly friendly face might conceal a plot, conspiracy, or coup, what is a rational response? Isn't it acceptable to lie to the liars? Most courtiers thought so. Even John of Salisbury said as much, claiming the virtuous few had to deceive on occasion to protect themselves from the evil schemers surrounding them. In *The Prince*, Niccolo Machiavelli (1513/1992) famously cautioned courtiers and monarchs alike to "never attempt to win by force what can be won by deception." Acknowledging that monarchs, unlike courtiers, must operate under the public expectation they will be virtuous and true, Machiavelli nonetheless promoted a more pragmatic mode of private counsel:

> Everyone admits how praiseworthy it is in a prince to keep his word, and to behave with integrity rather than cunning. Nevertheless our experience has been that those princes who have done great things have considered keeping their word of little account, and have known how to beguile men's minds by shrewdness and cunning. In the end, these princes have overcome those who have relied on keeping their word. (p. 69)

Does the same go for princesses? Christine de Pizan (1405/1999), a rare female courtier who served Charles VI of France, thought so. Royal women not only should always exude an air of honesty, she advised, but also must do their best to maintain harmonious relations with their husbands and other members of the court. When lying is the only way to achieve these goals, then lie they must.

These and other rationalizations of court chicanery reflect a true but troubling political insight—abiding by Augustine's absolutist doctrine

of truth-telling might win you a place in heaven, but lose your seat at the prince's table, lose the prince his kingdom, and lose the princess her prince. It's notable that the aforementioned courtly advisors rarely cite theologians (none mention Augustine!) in their recommendations, but all pay tribute to ancient Roman lawyer and rhetorician Marcus Tullius Cicero (Campbell, 2001). Cicero advised politicians to choose their words and actions based on circumstances, not immutable moral principles. In the circumstances of the royal courts, moral principles routinely conflicted with one another, such as the directives to be truthful at all times, to act with charity toward others, and to defend the reputation of one's allies. Courtiers often justified the "sin" of lying in one circumstance to avoid potentially worse transgressions in others (Denery, 2015).

This rhetorical framing of deception in royal courts complemented the probabilistic conception of truth emerging in criminal courts. Magistrates contended that, in the absence of reliable witnesses, the veracity of suspects can only be inferred to degrees less than absolute certainty, based on their demeanor and evidence pertaining to their circumstances. Courtiers maintained that the motives for deviating from veracity derive largely from circumstances—threat of punishment, promise of reward, pacts with allies, grudges against adversaries, etc. These ideas continue to drive the research emphasis on demeanor cues and motives in scholarly studies of deception in the modern era.

The Age of Enlightenment Through the Early Twentieth Century

The scientific revolutions beginning in the Enlightenment era of the seventeenth century and extending through the early decades of the twentieth century also profoundly shaped intellectual inquiry into deception. In some respects, this influence produced optimism. Educated people who embraced the "Enlightenment vision" could plausibly believe they could come to know important things about how the universe works (Searle, 1998). Seen through the lenses of the Copernican Revolution, Newtonian Mechanics, Maxwell's electromagnetic theory, and Darwin's theory of evolution, the contents of the world and universe seemed increasingly sensible and logical. It was also possible for scientific knowledge to peacefully coexist with religious belief, as long as one embraced Descartes' paradigmatic distinction between the physical/material and mental/spiritual realms. Even the "subversive" intellectual revolutionaries of the late nineteenth and early twentieth centuries, Karl Marx and Sigmund Freud, considered their scholarship to be contributions to Enlightenment-inspired scientific progress. Marx purported to be creating a science of history, Freud a science of mind.

Freud, his student Carl Jung, William James, and other psychology pioneers inspired some early forensic scientists to believe the probabilistic

assessment of suspect testimony could be replaced entirely by psychophysiological measurement (Trovillo, 1939). They reasoned that any effort on the part of a suspect to evade the truth could be circumvented by basing veracity judgments on biometric parameters such as blood pressure, pulse, heart rate, and galvanic skin response. Because these parameters are ostensibly involuntary, they were presumed counterfeit-proof evidence to corroborate or contradict a suspect's assertions. Italian physician and criminologist Cesare Lombroso (1895) was the first to translate this reasoning into instrumentation, which ultimately led to the creation of the polygraph two decades later. By the 1950s, over 2 million polygraph tests were being administered in the US every year by almost 10,000 examiners conducting police investigations or screening government and business personnel (Alder, 2007). Although few contemporary deception scholars consider the polygraph a highly reliable detection methodology (an issue we consider later), most continue to embrace its underlying logic that detection judgments benefit from a systematic evaluation of a speaker's verbal and nonverbal behaviors (Knapp et al., 2016).

Other ideas emerging during this period made intellectuals more pessimistic about their capacity to know themselves or their universe (Jameson, 1992; Tasic, 2001). Although Freud (1894, 1914) developed psychoanalysis as a method for people to learn about themselves and improve their mental health, it also implied that the self is in some respects unknowable, a tiny island of rational consciousness in a murky ocean of the irrational unconscious. Albert Einstein's (1916) relativity theory challenged fundamental assumptions about the relationships between space and time. According to this theory, if you travel the universe at the speed of light and return in 10 years, the world will be 90 years older than you are; how could that be? Quantum theorist Werner Heisenberg (1927) demonstrated that basic physical reality is indeterministic and efforts to observe it alter the reality being observed. Kurt Godel's (1931) incompleteness theorems demonstrated there are propositions in mathematical systems that count as "true" but cannot be proven to be true within those systems. Philosopher Ludwig Wittgenstein (1953) argued that human communication amounts to a series of mutually untranslatable "language games."

The relativistic implications of these ideas have been adapted and amplified in the contemporary intellectual movement known as "postmodernism." Postmodernists characteristically express doubt, distrust, and sometimes outright contempt for objective notions of truth and reality (Searle, 1998). The extent to which anything can be called "true" or "real," according to this view, derives entirely from a subjective individual or cultural perspective (Jameson, 1992). Despite drawing initial inspiration from scholars in math, physics, and psychology, postmodernism is primarily a school of thought in the humanities, with few proponents in the physical or behavioral sciences. The authors are not aware of any deception researchers who embrace this view, which makes sense—dismissing objective truth would make the task of

defining or detecting efforts to distort it rather difficult (McGlone, 2006). Nevertheless, the prevalence of postmodernism in cultural studies and popular culture has had a profound influence on contemporary society's cynicism about truth and attitudes about lying.

The Late Twentieth and Early Twenty-First Centuries

The previous section aptly demonstrates the words Shakespeare gave to one of his characters in *The Tempest*: "…what's past is prologue." Those older historical milestones set the stage for the more focused and intense examination of lying and deception beginning in the late 1960s in the US.

Public relations stunts were well established by 1961 when Daniel Boorstin wrote his book, *The Image: A Guide to Pseudo-Events in America*. But Boorstin felt that the variety of false and misleading events had reached a critical point—one in which the American culture was on the verge of creating a new reality for itself. Little did he know how far we would ride that pony. Pseudo-events are media vehicles that take the place of reality and are arranged for the sake of publicity or entertainment. For example, a movie studio leaks a story that a popular actor has died. Once that news is widely publicized, the studio calls a news conference to deny the death and, while they're at it, blame the false story on a competitor and announce a new movie with the reportedly fallen star. Boorstin maintained that

> …the American citizen lives in a world where fantasy is more real than reality, where the image has more dignity than its original…The pseudo-events that flood our consciousness are neither true nor false in the old familiar senses. The very same advances which have made them possible have also made the images—however planned, contrived, or distorted—more vivid, more attractive, more impressive, and more persuasive than reality itself. (p. 37)

While Boorstin infrequently used the words "lying" or "deception," the concept of pseudo-events clearly stands as a representative example of those categories and served as a forerunner for the many versions of lying and deception to follow.

The theme of creating false realities initiated by Boorstin was subsequently taken up by numerous authors and researchers. They examined it from a variety of perspectives. Each decade following the 1960s has seen more published books about lying and deception than the previous one. Scholarly publications in research journals have followed a similar pattern. Scholars from virtually every academic discipline have contributed books on the subject—anthropology (Bailey, 1991); art (Gombrich, 2000; Honeycutt, 2014); biology (Fujinami & Cunningham, 2000; Oldstone, 2005); botany (Alcock, 2005); communication (Knapp et al., 2016; Levine, 2014; Richards, 1990); economics (Akerlof & Shiller, 2015); entomology (Lloyd, 1986); history (Fernández-Armesto, 1997); journalism (Campbell, 2017; Paterno, 1997);

law (Perlmutter, 1998); management (Kihn, 2005); mathematics and statistics (Mauro, 1992; Seife, 2010); media studies (Mitchell, 1992); medicine and psychiatry (Dubovsky, 1997; Ford, 1996; Kucharski, 2014); philosophy (Nyberg, 1993); physics (Park, 2000); psychology (Ekman, 2001); political science and government (Campbell, 2017; Cliffe, Ramsay, & Bartlett, 2000; Paterno, 1997); public policy (Pfiffner, 2003); public relations and advertising (Boush, Friestad, & Wright, 2009; Richards, 1990); religion (Denery, 2015); sociology (Barnes, 1994); and zoology (Cloudsley-Thompson, 1980; Stevens, 2015). Since the late 1990s, there have also been a steadily increasing number of college courses devoted entirely to the subject of lying and deception and many more courses with units that explore this topic.

The increasing attention given to deceptive practices led some authors in the first decade of the twenty-first century to declare we had entered a new era. Keyes (2004) called it the "post-truth era"—a time when the lines between truth and lies, honesty and dishonesty, fiction and nonfiction were thoroughly blurred. Deceiving others, he argued, had become habitual. Keyes wasn't alone. In 2005, comedian Stephen Colbert introduced the term "truthiness" to describe what he believed was an all too common tendency for people to claim as truth something that they only knew intuitively or because it "feels right" rather than something based on facts, evidence, and/or reasoning. In that same year, Frankfurt (2005) and Penny (2005) depicted America as a society where "bullshit" was rampant. According to Penny, "we live in an era of unprecedented bullshit production" (p. 1). "Our era is unique by virtue of its sheer scale, its massive budget, its seemingly unlimited capability to send bullshit hurtling rapidly over the globe… Even a cursory study of bullshit yields an embarrassment of riches, an all-you-can-eat buffet of phoniness…" (p. 2). "I am even tempted to make the case that lying is less dangerous than bullshitting…The liar still cares about the truth. The bullshitter is unburdened by such concerns…Bullshit is forever putting the rosiest of spins on rotten political and economic decisions" (pp. 4–5). Bullshit, according to Frankfurt, "is unavoidable whenever circumstances require someone to talk without knowing what he is talking about" (p. 63).

Manjoo (2008), guided by these same perceptions, described what he called a "post-fact era"—a society overrun with the tendency to believe whatever outlandish thing you wanted to believe without any regard for facts. It is exemplified by politicians who endlessly repeat talking points while ignoring factual rebuttals. While most people who behave this way are reluctant to admit their disdain for facts, Jeffrey Lord, CNN analyst and former Reagan associate political director, had no reservations when he said: "… I honestly don't think this fact-checking business — as we're all into this — is anything more than, you know, one more sort of out-of-touch, elitist, media-type thing. I don't think people out here in America care. What they care about are what the candidates say" (Borchers, 2016). This was simply an echo of

what Neil Newhouse, a Romney pollster, said the week of the Republican National Convention in 2012: "…we're not going to let our campaign be dictated by fact-checkers" (Stein, 2012).

Given the belief that we had indeed entered a post-truth or post-fact era replete with bullshit and what Jackson and Jamieson (2007) called a "world of disinformation," it was not surprising to see a rise in the number of efforts to prevent lies and to publicize facts in response to undocumented claims and assertions confronting the public. One of the first groups to undertake the task of identifying and disseminating instances of public deception was the National Council of Teachers of English. During the Watergate scandal in 1971, the NCTE established the Committee on Public Doublespeak. Its purpose was to analyze, record, and publicize the way public officials, advertisers, and others use language to distort, mislead, and manipulate. The organization's *Doublespeak Award*, first presented since 1974, is given to a public spokesperson or advertisement in which the language is grossly deceptive, confusing, or evasive. Lutz (1989) argued the more disturbing linguistic distortions of reality are used by people in power to mislead others for their own purposes and, if not exposed, will structure the way we construe and experience reality.

In 1995, the Internet site *Snopes.com* was established to track down and clarify rumors, scams, urban legends, and other stories of unknown or questionable origin circulating on the Internet. In 2003, the Annenberg Public Policy Center at the University of Pennsylvania's Annenberg School of Communication launched *FactCheck*. *FactCheck* focuses primarily on political rhetoric and seeks facts to determine the validity of statements made by political candidates, officeholders, and other public officials. Another organization with similar aims, called *PolitiFact*, was established in 2007 and is operated by the *Tampa Bay Times*. Since 2009, they have publicized what they believe to be the "lie of the year." The goals of *WikiLeaks*, founded in 2006, are not to expose the deception in what is said as much as it is to expose secret and/or classified information that might be the basis of what is being said (or not said). They do not identify the names of their sources nor the means by which they obtained their information. While there are individuals who, because of their "insider" status, try to expose the lies of corporations and public entities, these "whistleblowers" do not always fare well if they are identified and therefore play a limited role in the efforts to expose those who deceive the public (Alford, 2001; Glazer & Glazer, 1989).

Some efforts have been made at the federal level to protect the consumer from deceptive messages, but free speech rights, poorly written laws, the difficulty in proving intent, and the lack of enforcement personnel often neutralize the effects of such legislation. Libel laws, designed to protect a person from lies that damage his or her reputation, have been in existence as long as the country itself. The *Fair Packaging and Labeling Act* of 1968 was designed to encourage honesty in product labeling. *The Truth in Lending Act*

of 1968 was aimed at eliminating deceptive practices related to the costs and terms associated with borrowing money. One of the jobs of the Federal Trade Commission is to protect the public from deceptive advertising. Recently, Congress introduced a bill that was intended to protect the public from photo-shopped images. In 2014 and again in 2016, a *Truth in Advertising Act* was proposed that would give the FTC the power to examine potential harm arising out of any media images in ads that were altered to materially change the appearance and physical characteristics of models' faces, bodies, skin color, weight, signs of aging, etc. Needless to say, the problems with enforcement alone are likely to doom this bill. But it is significant in that it is an effort to deal with a potentially deceptive tool that is unique to the twenty-first century and available to virtually anyone with a computer.

How did we get to this point? What led to the belief that we are being overrun with bullshit, living in a post-truth society, and badly in need of fact-checking organizations as a counterbalance? Like other subjects that dominate the public and academic mind-set, our current concern with lying and deception was fertilized and grew out of various social, political, and technological forces during the past half century.

Highly Publicized Incidents of Lying

Most people tell the truth most of the time. The maintenance of social cohesion demands it. But some people lie a lot and some lies are more visible and affect more people. These lies, often by public figures in positions of power, are responsible for creating a widespread awareness of deception and its effects. Examples of this during our recent past are plentiful.

That US presidents and aspirants to the office have engaged in willful deception is well documented (Alterman, 2004; Pfiffner, 1999). The vast volume of confirmed presidential lies prohibits detailed documentation here, but every past president probably lied at some point during his tenure. However, the Oval Office occupant at the time of this writing seems to have surpassed the deceit of his predecessors in both quantity and audacity. President Donald Trump's track record of untruths after just six months in office was a scandal of epic proportions. *New York Times* reporters David Leonhardt and Stuart Thompson (2017) artfully expressed the shock and dismay many White House observers felt while monitoring Trump's conduct:

> There is simply no precedent for an American president to spend so much time telling untruths. Every president has shaded the truth or told occasional whoppers. No other president — of either party — has behaved as Trump is behaving. He is trying to create an atmosphere in which reality is irrelevant.

These authors draw this depressing conclusion after documenting a "remarkable" feat Trump achieved at the very start of his term: He said something

untrue in a public statement every day of his first 40 days in office (June 20 through March 1, 2017).

There is no federal law that prevents politicians from lying in public statements or advertisements, and while some states have laws prohibiting political lies, free speech rights and the difficulty in proving intent make them very hard to enforce. So as Rue (1994, p. 246) pointed out: "There are many honest and truthful ways to elicit positive responses from voters, but it has long been recognized that they are less effective than deceptive means. Exaggeration, distortion, quoting out of context, innuendo, false promises, pandering, scare tactics, and flat-out-lies have become the standard fare of political campaigns." As a result, Miller and Stiff (1993) argued:

> …many citizens are becoming more permissive, or at least more fatalistically accepting, of deceptive tactics. Certainly cynicism about the veracity of politicians is a venerable characteristic of the American voter, but this cynicism has typically been coupled with belief in the moral culpability and responsibility of the offending party. During recent campaigns, many political commentators and voters alike seem to have become resigned to the fact that deceptive communication is merely part of the 'getting elected' game. This tone of resignation surfaces in statements justifying deceit on the grounds that "it was just something that was said during the campaign," the implication being that campaign pledges can be expected to become inoperative on inauguration day. To the extent that citizens accept a shift from a norm of honesty to a norm of deceit, traditional democratic values relating to the need for an informed populace and debate about the substance of issues will be seriously threatened. (p. 5)

Politicians haven't been the only ones producing deception for public consumption during the last half of the twentieth century and the first part of the twenty-first century. During the 1980s, the testimony of children about incredibly fanciful, dangerous, and unhealthy events at the McMartin day care grabbed national headlines. After many years, the legal system found these stories to be false, but this case and others like it prompted researchers to look closely at the lying and truth-telling behavior of children (Ceci & Bruck, 1995; Eberle & Eberle, 1993). Memoirs are life stories and are expected to be truthful. Sometimes memoirists are guilty of a faulty memory, but outright lies characterized several nationally well-known memoirs during this period (Frey, 2003; Menchú, 1983; Rosenblat, 2009; Wilkomirski, 1996). Professional journalists who made up stories, invented sources, quotes, and events, combined elements of several stories, or plagiarized stories also received national media exposure. Reporters from the *Washington Post*, *USA Today*, *The New York Times*, and the *New Republic*, among others, were discredited. Even Brian Williams, a respected broadcast journalist for NBC news, publicly lied in 2014 about his experiences in a combat zone (Farhi, 2015). And when the Governor of New Mexico, the Notre Dame football coach, the Poet Laureate of California, executives from Oracle, Radio Shack, Bausch & Lomb,

the US Olympic Committee, among others, fabricated information on their resumes, the public took notice. In 2013, Congress passed the *Stolen Valor Act* to prevent and/or punish the increasing number of people who were falsely claiming military service or medals.

Popular Books, TV Shows, and Movies

Self-help books like Liberman's (1998) *Never Be Lied to Again* and best sellers like Harvard philosopher Sissela Bok's (1978) *Lying: Moral Choice in Public and Private Life* helped to illumine the subject of deception for the masses. Former Secretary of Education William Bennett's edited *Book of Virtues* (1993), also a best seller, was followed by his *Children's Book of Virtues* (1995). Both of these books include honesty as one of the virtues to be admired and practiced, with lying as the villain. An animated television series based on the book for children ran on PBS from 1996 to 2000.

Beginning in the late 1960s, CBS Television's *60 Minutes* built a reputation with investigative journalism primarily focused on the lies of real individuals and corporations. The interest in lying was also a big part of television entertainment. As of 2018, the game show *To Tell the Truth* is one of only two game shows to have aired at least one new episode in each of the past seven decades. Some TV series have devoted a single episode of a drama or comedy to the subject of lying, but in recent years it has become the theme around which entire series are based. From 2009 to 2011, the show *Lie to Me* featured a character loosely based on researcher Paul Ekman who could detect lies by carefully observing a person's nonverbal behavior. Showtime's dark comedy series *House of Lies* (2012–2016) was based on Martin Kihn's (2005) account of the manipulative practices he witnessed and performed as a management consultant for *Booz Allen Hamilton*. ABC's mystery detective series *Secrets and Lies* was launched in 2015. A movie by the same name was released in 1996.

Disney's animated film, *Pinocchio*, was a critical success but a box office disaster in 1940. But in 1994, the Library of Congress admitted it into the National Film Registry. This tale of a boy whose nose grows when he lies has become a cultural icon—inspiring dozens of books, movies, and sculptures. The biggest box office success in March in movie history was the 1997 film, *Liar, Liar*, a comedy about a lawyer who was restrained from lying for 24 hours. *Catch Me If You Can*, a film based on the exploits of 1960s con artist Frank Abagnale, was both a critical and box office success in 2002. Ricky Gervais' (2009) movie *The Invention of Lying* tells the story of a liar living in an alternate universe in which it has never occurred to anyone else to twist the truth.

These books, TV shows, and movies represent only a fraction of the entertainment media attention given to lying and deception during this period.

COMPUTER-MEDIATED COMMUNICATION

Starting in 1991, the following became available to most of the American public over the next 20 years: the Internet, email, mobile phones, photo-editing software (Adobe *Photoshop, Corel Paintshop*, etc.), social media (Facebook, Twitter, etc.), and online dating sites (*eHarmony*, *OkCupid*, etc.). This revolution in communication technology brought with it the ability of an individual liar (or group) to spread lies faster and to more people than ever before—sometimes anonymously. Corporations, whose reputation and profits could suffer from Internet disinformation, often employ people to monitor blogs, newsgroups, and other sites for what is being said about them so they can act quickly to counteract it.

When *Photoshop* was made available to the general public, virtually anyone with a computer could alter a visual image. Sometimes these alterations were done for purely aesthetic or comedic purposes, but a growing number are produced with the express goals of damaging someone's reputation (cyberbullying, revenge porn), misleading others about one's appearance and accomplishments, or reinventing a discredited idea (like Bigfoot or aliens). Software improvements and the increasing skill of users often make it difficult to identify faked visual images. This dilemma has spawned a number of books on visual literacy (Barry, 1997; Brugioni, 1999; Messaris 1994; Mitchell, 1992) with some arguing that the greater number of faked images increases the likelihood that true images are ignored or questioned.

Computer and mobile phone applications now also provide a mechanism for users to shut themselves off from any news (customized "My News" apps) or friends whose beliefs differ from their own (defriending on Facebook). Thus, users can select and/or manufacture their own evidence. These applications provide welcoming outlets for disinformation while simultaneously cementing the divisions between those whose beliefs differ. Scientist and Internet critic David Helfand (2016) described it this way:

> Today, the climate change denier, homeopathic practitioner, or presidential candidate can easily, quickly, and cheaply raise armies of the uninformed, the gullible, and the disenchanted by providing their echo chambers with an endless diet of self-reinforcing nonsense. (p. 56)

Many disenchanted Americans gorged themselves on this diet during a wave of "fake news" during the presidential campaign of 2016 (Allcott & Gentzkow, 2017). Although false news stories designed for political ends are nothing. In the first century, Octavian famously dispatched couriers spreading false stories about rival Marc Antony to take control of the Roman republic; in the eighth century, the Catholic Church forged a Roman imperial decree alleging that 400 years earlier Emperor Constantine the Great had transferred extensive land and political power to Pope Sylvester I in exchange

for curing his leprosy; in 1782, Founding Father Ben Franklin drummed up European sympathy for the American revolutionary cause with a false news story purporting the British had hired Native Americans to scalp the colonists (D'Costa, 2016). But distributing these false stories via the social media platforms *Facebook* and *Twitter* enables the lies to spread faster, farther, and more frequently than ever before. In addition, the availability of online publishing platforms like *WordPress* allows fakers to create professional-looking digital distribution sites for the stories with ease.

Technology certainly played a key role in the proliferation of fake news in the Presidential election of 2016, especially with Donald Trump's often unsupported claims that *CNN*, *The New York Times*, *The Washington Post*, and other news media dissected and labeled false. These constant, but often critical, accounts of their candidate's claims repelled Trump supporters from mainstream news outlets. In an atmosphere where they didn't trust news stories critical of their candidate but also couldn't be certain what his next outlandish and off-the-cuff pronouncement might be, they became easy targets for the distortions and hyperbole characteristic of fake news (Priest & Birnbaum, 2017).

"Pope backs Trump," "Hillary sold weapons to ISIS," "FBI Agent Suspected in Hillary Email Leaks Found Dead"—these fake headlines all went viral on Facebook in the run up to the election, gaining such high engagement that *Buzzfeed* published analysis of how they had outperformed real news in terms of story posts, shares, comments, and reactions in *Facebook* (Silverman, 2016). Some of these stories appear to have been created by overzealous Trump supporters and/or Clinton detractors in the US; others were designed as simply "clickbait" to generate online advertising traffic, with no political motive. But the most prolific source of fake news appears to have been the Kremlin, which has a long history of election-meddling in the US and other countries. The goal of their state-of-the-art propaganda campaign was to punish candidate Hillary Clinton, help Donald Trump, and undermine faith in American democracy.

The effort also sought to heighten the appearance of international tensions and promote fear of looming hostilities with nuclear-armed Russia. The sophistication of the Russian tactics may complicate efforts by Facebook and Google to crack down on "fake news," as they have vowed to do after widespread complaints about the problem (Allcott & Gentzkow, 2017; Timberg, 2016).

If we are to withstand fake news epidemics in the future, we will have to sharpen our sense of skepticism and ask pertinent questions about the veracity of what we view and share. At a time when 60% of Americans get their news primarily through social media (Silverman, 2016), spreading propaganda requires only some Web space and a receptive audience willing to share it with their online communities.

Personal Relationships, Communication, and Nonverbal Behavior

The number of messages by public officials on social and political issues during the 1960s that were perceived by the public as intentionally secretive or outright manipulative not only led to a distrust of the people delivering the messages, but the mass media over which these messages were delivered. The social unrest that was generated probably reached its height during the Vietnam War. As a result, there was a widespread yearning for a more transparent society where message truth was more reliable and where the quality of one's life was anchored more in personal relationships.

Given the perceived lack of reliability associated with verbal messages emanating from the mass media, many believed that nonverbal signals would be a less-manipulated source of information. Nonverbal behavior, it was believed, was performed with little or no awareness by the communicator and therefore non-manipulative. It was hoped that these subtle cues might reveal unspoken prejudices and deceptive intent. Thus, learning how to "read" a person's behavior that they presumably had little or no control over seemed like a desirable skill to acquire. This was when the academic community initiated more studies designed to illumine the nature of nonverbal behavior and the role it played in lying and deception.

The academic interest in nonverbal behavior mirrored the growth of interest in studying communication in personal relationships. Sensitivity or encounter groups were fashionable in the 1970s. Participants in these small groups gained insights into the nuances of their interaction with others via unrestrained feedback and other techniques. The pros and cons of openness and "total honesty" were a common theme and prompted some to mistakenly think that "letting it all hang out" or what later came to be known as "radical honesty" (Blanton, 2005) was the secret to quality relationships at a time when the divorce rate was about 50%.

In the field of Communication Studies, the study of interpersonal communication was coming of age in the 1970s and two areas central to an understanding of honesty and deception were dominating the scholarship—self-disclosure and the credibility of speakers delivering persuasive messages. By the 1980s, the academic study of personal relationships was well established in college courses and the *Journal of Social and Personal Relationships* was launched in 1984. Another academic journal, *Personal Relationships*, followed in 1994 and two books specifically focused on deception in romantic relationships appeared at the turn of the century (Campbell, 2000; Forward, 1999).

The aforementioned societal activity set the stage for the growing interest and research on lying and deception. Thus far, the research has been primarily focused on individual and interpersonal behavior in three primary areas:

liar behavior, liar detection, and theories about liars and lie detection. Scholarly efforts in each of these areas will be discussed, but in order to understand the history of this phenomenon, it is also important to understand the nature of the construct being tracked. Lying and deception is a multi-faceted construct with an abundance of forms. Some have been more prominent in the research agenda than others.

CONCEPTIONS OF DECEPTION

Most of us think we know what a lie is. But as Montaigne observed: "…the reverse of truth has a hundred thousand forms, and a field indefinite, without bound or limit" (Hazlitt, 1877, p. 40). It is true there is an abundance of linguistic formulations that are so often used in the construction of lies that they are like kinfolk. McGlone and Knapp (2010) labeled these constructions the "blood relatives" of deception. Some of these blood relatives provide a mechanism for deceit and do it in a way that makes it harder to classify the act as a lie without greater knowledge of the communication context and the communicator's intent.

These linguistic constructions that often not only go hand-in-hand with lies but also occur in non-deceptive behavior include exaggeration, evasion, indirectness, ambiguity, imprecision, half-truths, euphemisms, and many more. Messages that imply, but do not actually state alterations from perceived truth, were called "devious messages" by Bowers, Elliott, and Desmond (1977). Even accurate statements can be used in misleading ways.

In the cynical days of the Watergate scandal, Herzog (1973) predicted that the many ways language can be used to mask messages would continue to increase, grow in sophistication, and ultimately replace more direct forms of lying:

> America will be the first civilization to eliminate lies. Soon in America, the lie will be superfluous, unnecessary, and will be buried. The lie is not vanishing because it is being killed off, like some hapless species of wildlife. It is not disappearing because it was legislated out of existence, like a noxious fume, or because it has atrophied from lack of use. Clearly, lies cannot be regarded as victims of higher morality. The lie is a casualty of progress…The new device that is making the lie obsolete can be called the Fake Factor, or for those who require still more trenchant terminology, the B.S. Factor…this factor causes a subtle skewing of sense, a distortion of meaning, without ever becoming an actual lie. (p. 15)

In order to "become an actual lie," a determination of intent is necessary and Herzog is saying that there are many linguistic formulations and contexts designed to make that extremely difficult.

So practically, lies are defined by the way people perceive certain features of communicative acts in context. Thus, perceptions may differ. An omission or ambiguous phrase may be deemed a lie by one person and not another; deemed a lie in one context and not another. From this perspective, the definition of a lie is broad, fluid, and variable. However, the definition of a lie as studied by the academic research community is much narrower.

Whereas "known" or "ground" truth may not always be clear in everyday life, it is clear in experimental research designs. Operational definitions used by researchers studying lies often include the word "misleading," implying the existence of a known truth. Whereas in everyday life, misleading behavior may be perpetrated by those unaware of what they are doing (e.g., children or someone who is delusional), scholarly research is focused on adults who are well aware of what they are doing. Whereas one's intent to mislead is often ambiguous in everyday life, it is built into scholarly studies of deception. Lies are clearly deliberate in these situations. "Intentional deception" is a phrase commonly used by academic researchers to describe the phenomenon they are studying. Whereas in everyday life a liar may lie because the target makes it known that they want to be lied to, academic research focuses on targets who are not asking to be lied to.

Despite the aforementioned similarities in the type of lie studied, there are some differences. For example, some studies ask people to role-play lies, whereas others study situations where people make their own decision to lie; some studies tell participants that if they are caught lying there will be a very unpleasant punishment, whereas other studies avoid any consequences for the liar; in some studies, lies are generated as a result of dialogue with another person, and in others, the liar is talking to a video camera or writing in a diary (Knapp et al., 2016).

STUDYING THE BEHAVIOR OF LIARS

Attempts to identify behavior exhibited by liars have a long history, but the invention of the polygraph in the early twentieth century marks an important milestone in the modern era. The polygraph was based on three assumptions that continued to undergird the research on liar behavior into the twenty-first century, namely: (1) Lies are told by people who are afraid of being caught because they are lying about something significant; (2) liars exhibit arousal when they lie because they are anxious about being caught; and (3) there are behaviors liars will reliably exhibit that will reveal their deception, the most accurate being those outside their control. The polygraph measured blood pressure, pulse, respiration, and skin conductivity on the assumption that liars will show dramatic changes in these physiological behaviors during acts of deception (compared to their truthful responses). While the number of false positives associated with the polygraph has been sufficient to keep its results

from being used as evidence in court or taken seriously in academic research, its use by law enforcement is still widespread. Law enforcement, like many people around the world, still believes there is a behavior or behaviors that people will manifest when lying and these behaviors can be measured and/or observed (Bond & The Global Deception Research Team, 2006).

In the period following the launch of the polygraph, only isolated studies addressed the possible link of behavior to deception. For example, Berrien and Huntington (1943) concluded that pupil instability was associated with lying. It wasn't until the early 1970s, when the study of nonverbal behavior was rapidly growing, that researchers again looked at the possibility that certain behaviors may be linked to lying. This time, instead of hidden physiological behavior, the focus was on observable visual and auditory cues. Initially, nonverbal behavior, like the physiological behavior measured by the polygraph, was thought to be largely out of the liar's conscious control and therefore a reliable source of information about liar behavior. Even though this belief was greatly modified later, it served as a trigger for the upsurge of research on liar behavior. The early work of Ekman and Friesen (1972, 1974) was influential. They made a distinction between leakage clues (behaviors that mistakenly reveal the truth) and deception clues (behavior that suggests a person is lying without revealing the truth) and argued that deception-related behavior was most likely evident in the feet/legs, followed by the hands, with the face the least likely source.

The early studies of nonverbal behavior associated with deception were dominated by social psychologists, but in the mid-1970s a growing number of scholars from the field of communication also turned their attention to the behavior of liars (Miller & Stiff, 1993). Complementing the study of liar nonverbal behavior, Knapp, Hart, and Dennis (1974) examined the verbal behavior of liars—a focus that gained increased attention in the 1980s. Over 20 different types of verbal deception have been discussed in the research literature, and there have been several proposals for organizing them into a meaningful classification scheme (Gupta, Sakamoto, & Ortony, 2013). In communication studies, verbal deception classification schemes can be grouped into 3 classes. One group is predicated on Grice's (1975) analysis of cooperation in conversation, according to which communicators comply with an implicit contract to make their utterances as informative, accurate, relevant, and detailed as required, but not more. McCornack's (1992) information manipulation theory treats these four "maxims" as classificatory dimensions, claiming verbal deception is produced by violating one or more of them. A variant developed by Burgoon, Buller, Guerro, Afifi, and Feldman (1996) adds to this set a fifth dimension, "personalization," akin to verbal immediacy. A second group focuses less on what speakers say than what they intend, such as causing a hearer to believe a falsehood or cease believing a truth (lies of commission), or to allow her to continue believing a falsehood or continue without believing a truth (lies of omission). Bradac (1983) argued that the intersection between these intentions and different types of

beliefs (about the world, other beliefs, and the consequences of utterances) accounted for many verbal deception forms. Vincent and Castelfranchi (1979) examined the interplay of intentions, beliefs, and goals in "indirect lies" such as insinuation, in which the speaker misleads chiefly through implicature (e.g., saying "I hope Matt doesn't leave work early today," implying he has done so without saying so). Third, Hopper and Bell (1984) developed a typology of deception strategies based on people's perceptions of the similarity in meaning between words and idioms denoting deceptive practices (exaggerating, kidding, teasing, etc.). Multidimensional scaling analyses revealed 3 underlying dimensions of deception vocabulary (evaluation, detectability, and premeditation), from which they proposed a typology of six strategies: lies, false verbal statements intended to deceive; fictions, such as white lies and myths; playing, such as joking or kidding; crimes, such as conspiracies or entrapment; masks, such as hypocrisy and evasion; and "unlies," such as deceptive implications.

Periodically, there have been efforts to review the research on liar behavior and assess the state of our knowledge (DePaulo et al., 2003; Knapp & Comadena, 1979; Knapp et al., 2016; Levine, 2014; Vrij, 2008; Zuckerman, DePaulo, & Rosenthal, 1981). While there have been many studies, there have also been many conflicting results. For example, some studies associate eye gaze aversion with lying while others don't; some indicate liars talk more while others find they talk less; some studies indicate more hesitations by liars while others find fewer. The number of studies supporting some behaviors is greater than others, suggesting a greater likelihood of their appearance during deception (DePaulo et al., 2003; Vrij, 2008) and certain behaviors are more likely in certain contexts. But so far, Ekman's (2001) assessment of the search for liar behavior seems valid:

> There is no sign of deceit itself—no gesture, facial expression, or muscle twitch that in and of itself means that a person is lying. There are only clues that the person is poorly prepared and clues of emotions that don't fit the person's line. (p. 80)

As a result, conclusions about liar behavior are not behaviorally specific. Instead, broad categories of liar behavior are identified, allowing a specific behavior to be included or not. For example, liars may exhibit negative behavior and/or affect, tense and/or nervous behavior, actions that are less forthcoming or distancing, and the telling of less plausible stories (DePaulo et al., 2003; Knapp et al., 1974).

The vast majority of studies focused on verbal and nonverbal cues associated with lying rely on passive observers who do not interact with the target person they are observing and must base their judgments on the target's responses to a scripted set of questions. This is not how lying/truth-telling judgments are made in everyday life. Interestingly, observing a suspect being questioned can lead to enhanced source believability or a probing effect

(Levine & McCornack, 2001). But in law enforcement interviews questions are not overly scripted—as they often are in research studies—and there is prior knowledge or evidence which can be used to develop questioning relative to the lie. Any answers provided throughout the interview can be used in strategic follow-up questions. This type of active questioning reflects the methods of skilled lie catchers. Simply interacting and probing a source may result in lie detection accuracies that reflect those of passive observation studies (e.g., slightly above chance; Levine & McCornack, 2001). But the strategic use of evidence and prompting for diagnostically useful information can lead to high judgment scores in experts and laypersons (Levine, 2014). Levine (2015) summarizes the findings of studies which examine a variety of active questioning methods and shows that accuracy rates range from 69 to 100%. After decades of studies reporting rates just above chance, these new methods and findings are generating excitement in the deception research domain.

Studying Deception Detection

Is it possible to train people to be significantly better lie detectors? Probably, but we don't currently have enough data to know exactly how effective training can be. Some training programs have increased accuracy rates dramatically (deTurck et al., 1990; Frank & Feeley, 2003), but Vrij (2008) found the average rate for trained observers and untrained observers to be 57 and 54%, respectively. Few training programs achieve accuracy rates as high as 65%. However, relying on training methods which match judges' background and professional experience (Hurley, Anker, Frank, Matsumoto, & Hwang, 2014; Shaw, Porter, & ten Brinke, 2013) or familiarity with the lie situation (Reinhard, Sporer, Scharmach, & Marksteiner, 2011) is important in obtaining higher-accuracy results. Also, Blair, Levine, and Vasquez (2015) provided judges with multiple opportunities to engage in a deception detection task over a six-week time period. They showed increases in accuracy rates from 69 to 89%. But, changing the context of the task from a cheating scenario to a mock robbery resulted in lowered accuracy scores. This further supports the notion that, while training does increase accuracy of deception detection, lie detection training in one context may not generalize to another context.

Training programs are more likely to be successful to the extent they are based on what deception researchers have learned about lying behavior and from effective detectors. Some important lessons learned from their research that lend themselves to training programs are as follows:

- Trainees must be deprogrammed regarding discredited stereotypes about liar behavior (Bond & The Global Deception Research Team, 2006);
- their training materials should focus on a variety of different truth-tellers and liars (Vrij, 2008);

- the stakes for the liars should be moderate to high (Shaw et al., 2013; Whelan, Wagstaff, & Wheatcroft, 2015)
- they need to get immediate feedback on their performance and plenty of individual attention from trainers (Elaad, 2003);
- they need plenty of time and opportunities to practice (Vrij, Mann, Robbins, & Robinson, 2006);
- they should be encouraged to talk about their lie/truth decisions so trainers can use this information to guide future performance (Frank & Feeley, 2003);
- they should be taught the value of attending to both nonverbal and verbal behavior and exposed to different models of lie detection (Levine, 2015).

Conclusion

The authors have witnessed the growth of deception research from a modest academic specialty to a major, multidisciplinary problem area. However, we harbor no pretension of being able to take the stand and tell you the truth, the whole truth, and nothing but about the history of deception research. Mindful of Santayana's snub to accounts of the past with which we began this chapter, we know ours is a selective and defective review of the growth period we participated in, let alone the epochs and eras leading up to it. As for what's to come in this exciting and developing field, we take inspiration from the great philosopher again: "We must welcome the future, remembering that soon it will be the past" (p. 114). And when that happens, we will leave it to younger scholars to botch the job their own way.

References

Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for fools: The economics of manipulation and deception*. Princeton, NJ: Princeton University Press.

Alcock, J. (2005). *An enthusiasm for orchids: Sex and deception in plant evolution*. Oxford, UK: Oxford University Press.

Alder, K. (2007). *The lie detectors: The history of an American obsession*. New York: Free Press.

Alford, C. F. (2001). *Whistleblowers: Broken lives and organizational power*. Ithaca, NY: Cornell University Press.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives, 31,* 211–236.

Alterman, E. (2004). *When presidents lie: A history of official deception and its consequences*. New York: Viking.

Andrews, R. M. (1994). *Law, magistracy, and crime in old regime Paris, 1735–1789*. Cambridge, UK: Cambridge University Press.

Bailey, F. G. (1991). *The prevalence of deceit*. Ithaca, NY: Cornell University Press.

Barnes, J. A. (1994). *A pack of lies*. New York: Cambridge University Press.

Barry, A. M. S. (1997). *Visual intelligence: Perception, image, and manipulation in visual communication*. Albany, NY: SUNY Press.

Bennett, W. J. (Ed.). (1993). *Book of Virtues*. New York, NY: Simon & Schuster.

Bennett, W. J. (Ed.). (1995). *The children's book of virtues*. New York, NY: Simon & Schuster.

Berrien, F., & Huntington, D. (1943). An exploratory study of pupillary responses during deception. *Journal of Experimental Psychology, 32,* 443–449.

Blanton, B. (2005). *Radical honesty: How to transform your life by telling the truth* (Rev. ed.). New York: Dell.

Blair, J. P., Levine, T. R., & Vasquez, B. E. (2015). Producing deception detection expertise. *Policing: An International Journal of Police Strategies & Management, 38*, 71–85.

Bok, S. (1978). *Lying: Moral choice in public and private life*. New York: Pantheon Books.

Bond, C. F., & The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74.

Boorstin, D. J. (1992). *The image: A guide to pseudo-events in America*. New York: Vintage Books.

Borchers, C. (2016, June 27). Jeffrey Lord's absurd claim that media fact-checking is 'elitist.' *The Washington Post*. https://www.washingtonpost.com/news/the-fix/wp/2016/06/27/jeffrey-lords-absurd-claim-that-media-fact-checking-is-elitist/.

Boush, D. M., Friestad, M., & Wright, P. (2009). *Deception in the marketplace: The psychology of deceptive persuasion and consumer self-protection*. New York, NY: Routledge.

Bowers, J. W., Elliott, N. D., & Desmond, R. J. (1977). Exploiting pragmatic rules: Devious messages. *Human Communication Research, 3,* 235–242.

Bradac, J. J. (1983). The language of lovers, flovers, and friends: Communicating in social and personal relationships. *Journal of Language and Social Psychology, 2,* 141–162.

Brugioni, D. A. (1999). *Photo fakery*. Dulles, VA: Brassey's.

Burgoon, J. K., Buller, D. B., Guerrero, L. K., Afifi, W. A., & Feldman, C. M. (1996). Interpersonal deception: XII. Information management dimensions underlying deceptive and truthful messages. *Communication Monographs, 63,* 50–69.

Campbell, S. (2000). *Romantic deception: The six signs he's lying*. Holbrook, MA: Adams Media.

Campbell, J. (2001). *The liar's tale: A history of falsehood*. NY: Norton.

Campbell, W. J. (2017). *Getting it wrong: Debunking the greatest myths in American journalism*. Oakland, CA: University of California Press.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.

Chand, D. (1980). *The Yajurveda, Sanskrit text with English translation* (3rd ed). Delhi, India: VVRI Press.

Cliffe, L., Ramsay, M., & Bartlett, D. (2000). *The politics of lying: Implications for democracy*. New York, NY: St. Martin's Press.

Cloudsley-Thompson, J. L. (1980). *Tooth and claw: Defensive strategies in the animal world*. London: J.M. Dent & Sons.

D'Costa, C. (2016). Three historical examples of fake news. *Scientific American Blog Network*. https://blogs.scientificamerican.com/anthropology-in-practice/three-historical-examples-of-fake-news/.

de Pizan, C. (1405/1999). *The book of the city of ladies* (R. Brown-Grant, Trans.). London, UK: Penguin.

de Turck, M. A., Harzlak, J. J., Bodhorn, D. J., & Texter, L. A. (1990). The effects of training social perceivers to detect deception from behavioral cues. *Communication Quarterly, 38,* 1–11.

Denery, D. (2015). *The devil wins: A history of lying from the garden of Eden to the enlightenment.* Princeton, NJ: Princeton University Press.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118.

Dubovsky, S. L. (1997). *Mind-body deceptions: The psychosomatics of everyday life.* New York, NY: W. W. Norton & Co.

Eberle, P., & Eberle, S. (1993). *The abuse of innocence: The Mamertine preschool trial.* Amherst, NY: Prometheus Books.

Einstein, A. (1916). *Relativity: The special and general theory.* New York, NY: H. Holt and Company.

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage.* New York: Norton.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication, 22,* 353–374.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology, 29,* 288–298.

Elaad, E. (2003). Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology, 17,* 349–363.

Farhi, P. (2015, February 4). Brian Williams admits that his story of coming under fire while in Iraq was false. *The Washington Post.* https://www.washingtonpost.com/lifestyle/style/brian-williams-admits-that-his-story-of-coming-under-fire-while-in-iraq-was-false/2015/02/04/d7fe32d0-acc0-11e4-9c91-e9d2f9fde644_story.html?utm_term=.1bcf3ec6a214.

Fernández-Armesto, F. (1997). *Truth: A history and guide for the perplexed.* New York: St. Martin's Press.

Ford, C. V. (1996). *Lies! Lies!! Lies!!! The psychology of deceit.* Washington, DC: American Psychiatric Press.

Ford, E. B. (2006). Lie detection: Historical, neuropsychiatric and legal dimensions. *International Journal of Law and Psychiatry, 29,* 159–177.

Forward, S. (1999). *When your lover is a liar.* New York: HarperCollins.

Frank, M. G., & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research, 31,* 58–75.

Frankfurt, H. G. (2005). *On bullshit.* Princeton, NJ: Princeton University Press.

Freud, S. (1894). *The neuro-psychoses of defence. Standard Edition, 3,* 41–61.

Freud, S. (1914). *Psychopathology of everyday life* (A. A. Brill, Trans.). New York, NY: Macmillan.

Frey, J. (2003). *A million little pieces.* New York: Doubleday.

Fujinami, M., & Cunningham, R. S. (2000). *Molecular mimicry, microbes, and auto-immunity.* Washington, DC: American Society for Microbiology.

Glazer, M. P., & Glazer, P. M. (1989). *The whistleblowers.* NY: Basic Books.

Godel, K. (1931). Über formal unentscheidbare Sitzer der Principia Mathematica und overwater Systems, I. *Mineshaft for Mathematics und Physics, 38,* 173–198.

Gombrich, E. H. (2000). *Art and illusion: A study in the psychology of pictorial representation.* Princeton, NJ: Princeton University Press.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41–58). New York: Academic Press.

Gupta, S., Sakamoto, K., & Ortony, A. (2013). Telling it like it isn't: A comprehensive approach to analyzing verbal deception. In F. Palmieri, L. Tummolini, R. Falcone & M. Mycelia (Eds.), *The goals of cognition. Essays in honor of Cristiano Castellanies* (pp. 206–233). London, UK: College Publications.

Hazlitt, W. C. (1877). *The essays of Montaigne* (C. Cotton, Trans.). London, UK: Reeves & Turner.

Heisenberg, W. (1927). Über den anschaulichen Inhale der quantentheoretischen Kinematics und Mechanics. *Zeitschrift fur physic, 43,* 172–198.

Helfand, D. J. (2016, September/October). The better angels of our nature vs. the internet. *Skeptical Inquirer, 40,* 55–56.

Herzog, A. (1973). *The b.s. factor: The theory and technique of faking it in America.* New York: Penguin Books.

Honeycutt, B. (2014). *The art of deception: Illusions to challenge the eye and the mind.* Watertown, MA: Imagine.

Hopper, R., & Bell, R. A. (1984). Broadening the deception construct. *Quarterly Journal of Speech, 70,* 288–302.

Hurley, C. M., Anker, A. E., Frank, M. G., Matsumoto, D., & Hwang, H. C. (2014). Background factors predicting accuracy and improvement in micro-expression recognition. *Motivation and Emotion, 38,* 700–714.

Jackson, B., & Jamieson, K. H. (2007). *UnSpun: Finding facts in a world of disinformation.* New York: Random House.

Jameson, F. (1992). *Postmodernism, or the cultural logic of late capitalism.* Durham, NC: Duke University Press.

John of Salisbury. (1159/1990). *Politicraticus* (C. J. Nederman, Trans.). Cambridge, UK: Cambridge University Press.

Keyes, R. (2004). *The post-truth era: Dishonesty and deception in contemporary life.* New York, NY: St. Martin's Press.

Kihn, M. (2005). *House of lies: How management consultants steal your watch and then tell you the time.* New York: Grand Central Publishing.

Knapp, M. L., & Comadena, M. E. (1979). Telling it like it isn't: A review of theory and research on deceptive communication. *Human Communication Research, 5,* 270–285.

Knapp, M. L., Hart, R. P., & Dennis, H. S. (1974). An exploration of deception as a communication construct. *Human Communication Research, 1,* 15–29.

Knapp, M. L., McGlone, M. S., Griffin, D. G., & Earnest, W. (2016). *Lying and deception in human interaction* (2nd ed.). Dubuque, IA: Kendall Hunt.

Kucharski, A. J. (2014). Immunity's illusion. *Scientific American, 311,* 80–85.

Langbein, J. H. (1977). *Torture and the law of proof: Europe and England in the ancient regime.* Chicago, IL: The University of Chicago Press.

Leonhardt, D., & Thompson, S. A. (2017, June 23). Trump's lies. *The New York Times.* https://www.nytimes.com/interactive/2017/06/23/opinion/trumps-lies.html.

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392.

Levine, T. R. (2015). New and improved accuracy findings in deception detection research. *Current Opinion in Psychology, 6,* 1–5.

Levine, T. R., & McCornack, S. A. (2001). Behavioral adaptation, confidence and heuristic-based explanations of the probing effect. *Human Communication Research, 27,* 471–502.

Lloyd, J. E. (1986). Firefly communication and deception: "Oh, what a tangled web". In R. W. Mitchell & N. S. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 113–128). New York: State University of New York Press.

Lombroso, C. (1895). *L'homme criminel.* Paris, FR: Becca Freres.

Lutz, W. (1989). *Doublespeak.* New York: Harper & Row.

Macchiavelli, N. (1513/1992). *The prince* (N. H. Thomson, Trans.). Mineola, NY: Dover Publications.

Mackay, C. (1857). *Memoir of extraordinary popular delusions and the madness of crowds.* London, UK: Office of the National Illustrated Library.

Manjoo, F. (2008). *True enough.* Hoboken, NJ: Wiley.

Mauro, J. (1992). *Statistical deception at work.* New York, NY: Routledge.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs, 59,* 1–16.

McGlone, M. S. (2006). Contextomy: The art of quoting out of context. *Media, Culture and Society, 27,* 511–522.

McGlone, M. S., & Knapp, M. L. (Eds.). (2010). *The interplay of truth and deception.* New York, NY: Routledge.

Menchú, R. (1983). *I, Rigoberta Menchú: An Indian woman in Guatemala.* London, UK: Verso.

Messaris, P. (1994). *Visual literacy: Image, mind, and reality.* Boulder, CO: Westview.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication.* Newbury Park, CA: Sage.

Mitchell, W. J. (1992). *The reconfigured eye: Visual truth in the post-photographic era.* Cambridge, MA: MIT Press.

Muldowney, M. S. (2002). *Augustine's treatises on various subjects.* New York, NY: Catholic University of America Press.

Nyberg, D. (1993). *The varnished truth: Truth telling and deceiving in ordinary life.* Chicago: University of Chicago Press.

Oldstone, M. B. A. (2005). *Molecular mimicry.* New York, NY: Springer.

Park, R. (2000). *Voodoo science: The road from foolishness to fraud.* New York: Oxford University Press.

Paterno, S. (1997, May). The lying game. *American Journalism Review, 19,* 40–45.

Penny, L. (2005). *Your call is important to us: The truth about bullshit.* New York, NY: Crown.

Perlmutter, M. (1998). *Why lawyers (and the rest of us) lie and engage in other repugnant behavior.* Austin, TX: Bright Books.

Pfiffner, J. P. (1999). The contemporary presidency: Presidential lies. *Presidential Studies Quarterly, 29,* 903–917.

Pfiffner, J. P. (2003). *The character factor: How we judge America's presidents.* College Station, TX: Texas A & M Press.

Priest, D., & Birnbaum, S. (2017, June 25). Europe has been working hard to expose Russian meddling for years. *The Washington Post.* https://www.washingtonpost.com/world/europe/europe-has-been-working-to-expose-russian-meddling-for-years/2017/06/25/e42dcece-4a09-11e7-9669-250d0b15f83b_story.html?noredirect=on&utm_term=.bc723421ff61.

Reinhard, M. A., Sporer, S. L., Scharmach, M., & Marksteiner, T. (2011). Listening, not watching: Situational familiarity and the ability to detect deception. *Journal of Personality and Social Psychology, 101,* 467–484.

Richards, J. I. (1990). *Deceptive advertising*. Hillsdale, NJ: Lawrence Erlbaum.

Rosenblat, H. (2009). *Angel at the fence: The true story of a love that survived*. New York, NY: Berkley Books.

Rue, L. (1994). *By the grace of guile: The role of deception in natural history and human affairs*. Oxford, UK: Oxford University Press.

Santayana, G. (1906/1998). *The life of reason*. Amherst, NY: Prometheus Books.

Searle, J. R. (1998). *Mind, language, and society*. New York, NY: Basic Books.

Seife, C. (2010). *Proofiness: The dark arts of mathematical deception*. New York, NY: Penguin.

Shaw, J., Porter, S., & ten Brinke, L. (2013). Catching liars: Training mental health and legal professionals to detect high-stakes lies. *The Journal of Forensic Psychiatry and Psychology, 24,* 145–159.

Silverman, C. (2016, November 16). This analysis shows how viral fake election news stories outperformed real news on Facebook. https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook.

Stein, S. (2012, August 8). Mitt Romney campaign: We will not be 'Dictated by fact-checkers'. *The Huffington Post*. http://www.huffingtonpost.com/2012/08/23/mitt-romney-_n_1836139.html.

Stevens, M. (2015). *Cheats and deceits: How plants and animals mislead*. Oxford, UK: Oxford University Press.

Tasic, V. (2001). *Mathematics and the roots of postmodern thought*. Oxford, UK: Oxford University Press.

Timberg, C. (2016, November 24). Russian propaganda effort helped spread 'fake news' during election, experts say. *The Washington Post*. https://www.washingtonpost.com/business/economy/russian-propaganda-effort-helped-spread-fake-news-during-election-experts-say/2016/11/24/793903b6-8a40-4ca9-b712-716af66098fe_story.html?utm_term=.bf009a97c450.

Trovillo, P. (1939). A history of lie detection. *American Journal of Police Science, 30,* 104–119.

Vincent, J. M., & Castelfranchi, C. (1979). On the art of deception: How to lie while saying the truth. In H. Parret, M. Sbisa, & J. Verschueren (Eds.), *Conference on pragmatics* (pp. 749–778). Trente, IT: Urbino.

Vrij, A. (2008). *Detecting lies and deceit: The psychology of lying and the implications for professional practice* (2nd ed.). Cheshire, UK: Wiley.

Vrij, A., Mann, S., Robbins, E., & Robinson, M. (2006). Police officers ability to detect deception in high stakes situations and in repeated lie detection tests. *Applied Cognitive Psychology, 20,* 741–755.

Whelan, C. W., Wagstaff, G. F., & Wheatcroft, J. M. (2015). Subjective to deception/honesty in a high-stakes situation: An exploratory approach. *The Journal of Psychology: Interdisciplinary and Applied, 149*, 517–534.

Wilkomirski, B. (1996). *Fragments: Memories of a wartime childhood*. New York, NY: Schocken Books.

Wittgenstein, L. (1953). *Philosophical investigations*. London, UK: Blackwell.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). New York: Academic Press.

CHAPTER 2

Defining Truthfulness, Deception, and Related Concepts

Pamela J. Kalbfleisch and Tony Docan-Morgan

What constitutes truthfulness and deception in everyday human interaction is not a simple, straightforward phenomenon. Further, truthfulness, deception, and the gray areas in-between cut across nearly all contexts of communication and relationship types (e.g., Carter, 2014; Docan-Morgan, 2007, 2011; Docan-Morgan & Manusov, 2009; Kalbfleisch, 1990, 1994; Mazur & Kalbfleisch, 2003; Stearns, 2014b). In the current chapter, we provide readers with a concise understanding of truthfulness, deception, and related concepts. We also raise important questions about the competing desires humans have regarding truthful and deceptive communication, various gray areas where the nature of truth and deception are blurred, and the ethics of potentially deceptive messages.

Truthful and Deceptive Communication: A Concise Overview

Truthfulness, as well as people's perceptions of it, surrounds much of our everyday communication. In most contexts of human interaction, truth-telling is expected to be the norm, while deception is the aberration.

P. J. Kalbfleisch (✉)
Department of Communication, University of North Dakota, Grand Forks, ND, USA
e-mail: pamela.kalbfleisch@email.und.edu

T. Docan-Morgan
Department of Communication Studies, University of Wisconsin-La Crosse, La Crosse, WI, USA
e-mail: tdocan@uwlax.edu

29

Communicating with *truthfulness* involves the exchange of information that is known by the communicator to be accurate (i.e., the accurate portrayal of information as understood by the sender). In other words, to tell the truth is to communicate with fidelity. The truth can be told verbally or nonverbally, through words, or by actions.

Communicating with truthfulness, however, is not necessarily a straightforward phenomenon—what constitutes the truth may vary from one individual to another, is based on perception and understanding of information, can change quickly or over time, and is often surrounded by innumerable contextual factors. Although the notion of truth is nebulous, humans tend to believe that they are usually being told the truth. Levine's (2014a) truth-default theory explores these issues in more depth:

> The central idea behind truth-default theory is that people tend to presume that other people communicate honestly most of the time. The presumption of honesty enables efficient communication and cooperation. Furthermore, since most people are honest most of time, believing others usually results in correct belief states. (p. 390)

Most of us do not cast continual doubt about the veracity of every person's messages. For a moment, imagine how messy human interaction would be if we constantly had to determine the truthfulness of every message. Of course, there are particular individuals whom we might doubt perpetually (e.g., an ex-spouse who has a track record in lying, stealing, and infidelity), and others whose job requires them to cast doubt on others (e.g., criminal interrogators). However, these are the exceptions to the norm. Most of us do not approach interactions with suspicion asking questions such as: "Are they deceiving or telling the truth? Am I being lied to? What are they lying about?"

Communicating with someone who we believe is telling the truth, for some, is like wearing an old pair of shoes. The Danish use the word *hygge*, roughly translated as comfort, connected, snug, and secure (Wilking, 2016). This comfortable and trusting state is easily violated when one finds they are the receiver of deceptive communication. We can safely assume that human beings seek comfortable environments and avoid discomfort; likewise, regarding communication, individuals naturally expect most of the messages they are told to be truthful. Indeed, in the vast majority of interactions, no one wants to be the recipient of deception. Additionally, it is simply easier and less taxing to assume others are telling the truth. However, a conundrum occurs when one perceives or discovers that they are being deceived. Compared to deceptive communication, truthful communication generally meets our expectations, helps us move forward in our interactions and understanding of one another, and may bring a level of ease to the interaction. Of course, not every truthful message is pleasant; yet, the assumption that our interaction partner is telling the truth, in contrast to lying, is much more likely to engender a sense of trust, comfort, and security.

While truthfulness involves communicating with fidelity (i.e., conveying a message that is known by the communicator to be accurate), deception involves the converse—a violation in the assumed fidelity of communication. *Deception* has been defined in a multitude of ways and for various purposes (see Chapter 1 by McGlone and Knapp and Chapter 14 by Carr, Solbu, and Frank). We provide an overview of modern definitions and explanations of *deception* and *lying* from notable authors in numerous fields of study. The list we provide is not exhaustive; instead, our goal is to share with readers various explanations, allowing us to comprehend similarities and differences, as well as key complexities when trying to define, understand, and study deception and lying.

1. Sissela Bok, a philosopher and ethicist (Stearns, 2014a), defines *deception* and *lying* early on in her classic book, *Lying: Moral Choice in Public and Private Life* (1978):

   When we undertake to deceive others intentionally, we communicate messages meant to mislead them, meant to make them believe what we ourselves do not believe. We can do so through gesture, through disguise, by means of action or inaction, even through silence. Which of these innumerable deceptive messages are also lies? I shall define as a lie any intentionally deceptive message which is *stated*. Such statements are most often made verbally or in writing, but can of course also be conveyed via smoke signals, Morse code, sign language, and the like. Deception, then, is the larger category, and lying forms part of it. (p. 14)

2. Bella DePaulo, a social psychologist and renowned deception researcher (Sternglanz, Morris, & Makiyil, 2014), defines deception as "a deliberate attempt to mislead others" (DePaulo et al., 2003, p. 74).

3. Aldert Vrij, a professor in applied social psychology, as well as an internationally recognized expert in deception (Sandler, 2014), defines deception as "a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief which the communicator considers to be untrue" (Vrij, 2000, p. 6).

4. Mark Frank, a social psychologist by training, professor of communication, and "a well-known and respected deception researcher" (Levine, 2014b, p. 393), draws in part from his colleague Paul Ekman's work in the below discussion of deception and lying:

   Often the terms *deception* and *lying* are used interchangeably, but we think there is an important difference. We believe deception is the superordinate category, of which one subcategory is telling a lie. We define *deception* as any action or phenomenon that misleads someone; *lying* is an act whereby someone *deliberately* misleads another and does so without notifying that person that he or she will be misleading them (Ekman, 1985/2001). The words *deliberate* and *prior notification* are the crucial distinguishing

characteristics of a lie. Deception may or may not be a deliberate act, whereas a lie is always deliberate. (Frank & Svetieva, 2013, p. 115)

5. Timothy Levine, a communication scholar, internationally recognized leader in deception research, and author of truth-default theory, remarks, "Deception is defined as intentionally, knowingly, and/or purposely misleading another person" (Levine, 2014a, p. 379). He defines a lie as "a subtype of deception that involves outright falsehood, which is consciously known to be false by the teller, and is not signaled as false to the message recipient" (p. 380). Further, he articulates, "Other forms of deception include omission, evasion, equivocation, and generating false conclusions with objectively true information" (pp. 380–381).

6. Mark Knapp, whose research has made significant contributions to the areas of interpersonal and nonverbal communication, relationship development, and deception (Summary, 2014), highlights the complexities of defining lying and deception in both his sole-authored first edition of *Lying and Deception in Human Interaction* (Knapp, 2008) and co-authored second edition (Knapp, McGlone, Griffin, & Earnest, 2016). Knapp et al. (2016) remark that the only clear distinction between lying and deception "is that deception is normally considered a superordinate term that encompasses various fraudulent, tricky, and/or misleading behavior—including lies" (p. 10). Further, they state, "Practically, lies and deception are defined by the way people perceive certain features of communicative acts in context" (p. 11). Knapp et al. (2016) argue that more often than not, the extent to which we believe a person has or hasn't lied hinges on how people perceive the following five features: perceptions of awareness, perceptions of altering information, perceptions of intent, perceptions of the situation, and perceptions of effects or consequences. Further, they articulate the following:

Answering the question, 'What is a lie?' or 'What is deceit?' can best be done by finding out how people perceive various component parts of the transaction in question. Was the communicator aware of what he or she was doing? Did the communicator alter information he or she knew to be true? What was the intent or motive behind the communicator's message? Was there anything about this situation that would encourage lying or even authorize it? What consequences resulted from the communicator's behavior? Perceptions associated with these questions lead us to make attributions about whether lying or deception has occurred, whether it is serious or not, and to what extent it can or should be sanctioned. (p. 15)

There is clear support among scholars that deception is a superordinate or higher-level term that includes various behaviors, such as lying, omission, evasion, and equivocation. Upon a review of the definitions and explanations stated above, as well as an examination of the larger literature, human deception appears to include the following components: communicators,

information exchange, and knowingly misleading or inaccurate information. To synthesize previous definitions and provide a potentially useful conceptualization, we offer a parsimonious definition of deception that incorporates these elements. We define *deception* as the communication of information that is known by the communicator to be inaccurate and/or misleading. This communication can be verbal or nonverbal, in writing, or in actions. Deception can be enacted through manipulation of words, behaviors, or appearance so that others will form a false impression. Indeed, deceptive communication is multifarious; examples include bald-faced lies, confabulation, disinformation, fabrication, false denials, half-truths, high stakes lies, life saving lies, malingering, outright lying, withholding information, omission, puffery, spin, and strategic ambiguity. What constitutes deception in daily interactions, of course, is context dependent. We also echo Knapp et al.'s (2016) observation that deception is "defined by the way people perceive certain features of communicative acts in context" (p. 11). In other words, in everyday interactions, what constitutes deception depends on the set of circumstances that surround a particular situation or event, and is based on communicators' perceptions of various factors (e.g., intent, consequences).

The act of deception violates communicators' assumption of and expectation for truthfulness in everyday interaction (Grice, 1975; Levine, 2014a). However, our expectation for truth, as well as the fidelity of others' messages, can be violated in countless ways, whether through the communication of false information, creation of an erroneous impression, or through communication that allows an inaccuracy to go uncorrected and perhaps later discovered (see Kalbfleisch, 1992, 2001). When individuals perceive or discover deception, the perceived veracity of all previous and future communication can easily become suspect. Deceptive communication generally takes a toll on relational quality, regardless of the motivation of the deceit (e.g., to benefit the others, to benefit the self) (Kalbfleisch, 2001).

COMPETING DESIRES FOR TRUTHFUL AND DECEPTIVE COMMUNICATION

As comfort-seeking beings, communicators may find it easier in particular situations to deceive rather than to provide accurate information. It may simply be easier to tell a friend that they look professional in their suit, when in fact they appear dated and unstylish. Other options in this scenario might include *equivocation* (i.e., providing information that can be interpreted several different ways; e.g., "That suit looks interesting on you.") and *avoidance* (i.e., not providing information; e.g., "I'm hungry; let's get dinner."). These types of responses allow one to maintain comfortable interactions without directly providing honest, accurate information. Yet, they can each create a false reality in the receiver.

On the other hand, the truth may not be what one really wants to hear. One may prefer to think that they are the epitome of good health rather than hearing a poor medical report from their physician. People sometimes comfort themselves with the maxim that *no news is good news*, when in fact *no*

news is simply no news. Not hearing the results of a medical test may be more comfortable than hearing unfavorable results, but it is likely more useful to receive the truth rather than deceiving oneself *that no news is good news.* No news is sometimes perceived as "good," only because it may be more comfortable than hearing the truth.

Further, it can be difficult for communicators to know when they are being told the truth. It can be disquieting when an individual hears multiple conflicting perspectives from others. One may look for second and third opinions, read accounts of others in similar situations, try to review and understand research reports, and attempt to make sense of a confusing situation. The search for truth can turn into a discomforting situation, even making individuals vulnerable to others' ill-intended behaviors. For example, *confidence artists* (*a.k.a.*, con artists) are successful in gaining the trust of vulnerable individuals. They may offer an assured answer in a confusing situation, seem more reliable than others, and/or build a comfortable relationship with someone from whom they wish to gain money, prestige, power, or other desired ends. A con artist can take advantage of a vulnerable person's predisposition to believe that the other is being truthful. As the only or the most confident, safe, and secure voice, con artists can become the sole voice that vulnerable individuals trust (see Konnikova, 2016).

Across nearly all contexts of communication—interpersonal, small group, organizational, professional, and mediated interactions—we face moments where there are competing desires for truthful and deceptive communication. Should I tell the complete truth to my partner? Will my friend trust me in the future if I am caught? Is my boss being completely honest when she provides my performance appraisal? Do I really want to know the truth about my performance at work? As the leader of a high-powered entity, how will I be perceived if I use equivocation in my responses? The dialectical tensions or competing desires we have for truthful and deceptive communication can make human interaction a complicated dance. In these moments, we are pulled in opposing directions when making communicative choices, and our decisions to tell the truth or deceive may have short- or long-term consequences.

A Glance at Some Gray Areas of Deceptive Communication

Although people tend to presume that others communicate honestly most of the time, yet at times we experience competing desires between telling the truth and deceiving, it is the gray areas of deception—the *in-between* and often unknown—that offer rich opportunities to further conceptualize what constitutes truthful and deceptive behavior. Further, these gray areas help us begin to unravel or at least actively contemplate the various ethical components of potentially deceptive behavior. The below examples allow us to better understand some of the murky waters of truthful and deceptive communication.

One context where truthfulness and deception are present, yet may be considered gray in nature, is entertainment. For example, a magician uses props, distraction, and dexterity to create illusions intended to outwit the audience. Further, the audience at least tacitly agrees to being misled or tricked. Another particular gray area of truth and deception in the context of entertainment emerges in reality television programming. Although these shows are thought to be "reality" and expected to be authentic by many viewers, they are often highly scripted and edited in strategic ways. One publicized example comes from HGTV's *House Hunters*, a show that features a buyer or buyers touring numerous homes for sale and then purchasing one in the end. First-hand testimony from individuals on the show is particularly eye-opening. One participant, Bobi Jensen, mentioned that the show:

> …didn't even "accept" us being a subject for the show until we closed on the house we were buying. So then when they decided to film our episode we had to scramble to find houses to tour and pretend we were considering. The ones we looked at weren't even for sale…they were just our two friends' houses who were nice enough to madly clean for days in preparation for the cameras! (Hooked on Houses, 2012)

Entertainment Weekly, a magazine that focuses on entertainment media news and critical reviews, investigated the above claim. They received a statement from a publicist for *House Hunters*, which reads in part:

> We're making a television show, so we manage certain production and time constraints, while honoring the home buying process. To maximize production time, we seek out families who are pretty far along in the process. Often everything moves much more quickly than we can anticipate, so we go back and revisit some of the homes that the family has already seen and we capture their authentic reactions. Because the stakes in real estate are so high, these homeowners always find themselves RIGHT back in the moment, experiencing the same emotions and reactions to these properties. (Strecker, 2012)

Scholarly studies investigating viewers' perceptions of reality television programming have found in part that audiences expect these shows to be authentic or "real". In a study investigating *Survivor*, Crew (2006) discovered that viewers "generally trusted that the 'truth' of what they had saw had not been significantly altered" (p. 72). However, reality television has been deemed an exaggerated version of everyday life, and audiences and producers are forced to negotiate the gulf between representations and truth (Escoffery, 2006; see also Hall, 2006; Tsay-Vogel & Krakowiak, 2017). Yet, we wonder, at what point, if any, do false illusions of reality for the sake of entertainment become ethically questionable? If we believe that entertainment imparts values and has the potential to shape behavior, it seems important to consider the implications of deception in the context of reality television programming.

Another gray area of deception emerges in the context of strategic communication. Strategic communication is often used to describe how persuasion functions in applied settings such as advertising, public relations, and public affairs. Public relations professionals, for example, use strategic communication as they design internal and external communications to make a company's actions or company policies seem as desirable as possible. In some instances, they may utilize *spin* by crafting negative information positively (Kurylo, 2014). According to Knapp et al. (2016):

> Putting a spin on a story simply means that the communicator finds a way to make it look like something it isn't…The communicator's goal is to redirect the target's thinking in a way that is favorable to his or her point of view…The use of half-truths and refocusing techniques are two common ways to spin. (p. 190)

The use of spin, half-truths, and refocusing easily become gray in nature, and therefore innumerable questions arise: Are these strategies helping communicators engage in truthful, deceptive, and/or ethical behavior? Are professionals who utilize these techniques being deceptive, simply employing skilled strategic communication, or both? Further, is truthful strategic communication possible and/or desired, and what are its parameters?

Another context of where gray areas of truth and deception emerge is in advertising. The use of *native advertising*, or ads that mimic the media in which they are imbedded (Campbell & Marks, 2015), raises important considerations. Audiences may find themselves reading what appears to be "news stories," but in actuality are advertisements for services or products. These "stories" are sometimes labeled as "sponsored content," but can be easily overlooked by readers. Such content shared in a social media context, for example, comes with its own personalized endorsements from users' online communities (e.g., Facebook friends, Twitter and Instagram followers). Advertisers may insist that such methods are not deceptive, and justify the use of native advertisements via the fact that customers chose to read their ads. Native advertisements are cast as preferable to more traditional advertising strategies because customer media usage is not interrupted by more typical, classic ads. Native advertisements become a seamless part of the media experience, yet may also mislead individuals into thinking that they are traditional news stories. Therefore, we ask: Is native advertising truthful, deceptive, or somewhere in-between? Are these ads purposefully meant to mislead readers? What are the various ethical considerations of this type of advertising?

Truthful and deceptive communication, as well as the gray areas in-between, is also prevalent in the context of healthcare. These issues surface via patients who deceive physicians in order to receive unneeded healthcare services, healthcare professionals who provide erroneous information to patients, and deceptive pharmaceutical company claims and advertising (see Hubbell, 2014). In their discussion of drug advertising, for example, Pallegedara and Hancock (2014) address deceptive and gray-area issues:

> Depicting a family in a drug ad, even though the drug is only approved for adults aged 18 years and up, can be deceptive, insinuating that the product is safe to use for the whole family. Ads tend to use spokespersons who endorse the drugmaker's claims about a drug's efficacy. For example, drug manufacturer Pfizer used Dr. Robert Jarvik as the spokesman for the drug Lipitor. In the television ad, Jarvik was introduced as the inventor of the artificial heart, depicting him as an authority on cardiovascular issues and medical advice. The advertisement, however, failed to mention that Jarvik was not licensed to practice medicine. The advertisement depicted Jarvik as a user of Lipitor, though he later admitted to using Lipitor only after he became a spokesman. In addition, the ad showed Jarvik rowing a boat, when in fact the man rowing was a body double. By presenting Jarvik as a physically fit individual, the ad might have misled consumers into believing that a drug may have additional positive side effects. The Lipitor ad was pulled from circulation after these facts came to light. (p. 313)

The above example is noteworthy, as it points largely to instances of *potentially* misleading information—another consideration for gray areas of deception. Clearly, information that has the possibility of misleading an individual about a product that could affect their health is cause for alarm. The use of information that is potentially misleading raises various issues. For example, in what contexts or situations, if any, is the use of information that *may* mislead individuals ethically sound? How can we determine if a message has the potential to mislead individuals?

Our discussion of the above-mentioned contexts and examples—entertainment, strategic communication, advertising, and healthcare—offers a glance at some gray areas of truth and deception. As students, scholars, and practitioners, we must continue to grapple with what constitutes truth, deception, and the murky waters in-between, as well as the ethical boundaries of communicating these messages within and across contexts. As communication teachers and scholars, the authors of this chapter believe that ethical communication is fundamental to rewarding lives and healthy relationships, whether these relationships are interpersonal, intergroup, or international. At a time when debates surrounding truth and ethics are prominent in many realms of public discourse, students and scholars of this topic are presented with a fruitful opportunity to discuss and potentially shape the meaning and value of effective, ethical communication.

Conclusion

The intent of this chapter was to present readers with a concise overview of truthfulness and deception, as well as discuss various ways in which deception emerges in human interaction. We also explored the notion that humans assume that others communicate honestly most of the time, yet also experience competing desires between truthful and deceptive communication. Additionally, the gray areas of deception offer opportunities to further conceptualize what constitutes truthful, deceptive, and ethical behavior.

As important, we also raised significant questions about the aforementioned topics, as well as the ethics of potentially deceptive messages. We also presented numerous contexts in which deception has been and will likely continue to be practiced, all of which are fertile grounds for the continued study of truth-telling and deceit. The fascinating and complex topic of deceptive communication will captivate our minds ad-infinitum.

References

Bok, S. (1978). *Lying: Moral choice in public and private life*. New York: Vintage Books.

Campbell, C., & Marks, L. J. (2015). Good native advertising isn't a secret. *Business Horizons, 58,* 599–606.

Carter, E. (2014). Deception in different contexts. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 265–268). Thousand Oaks, CA: Sage.

Crew, R. E. (2006). Viewer interpretations of reality television: How real is *Survivor* for its viewers? In D. S. Escoffery (Ed.), *How real is reality TV? Essays on representation and truth* (pp. 61–77). Jefferson, NC: McFarland & Company.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118.

Docan-Morgan, T. (2007). Training law enforcement officers to detect deception: A critique of previous research and framework for the future. *Applied Psychology in Criminal Justice, 3,* 143–171.

Docan-Morgan, T. (2011). "Everything changed": Relational turning point events in college teacher-student relationships from teachers' perspectives. *Communication Education, 60,* 20–50.

Docan-Morgan, T., & Manusov, V. (2009). Relational turning point events and their outcomes in college teacher-student relationships from students' perspectives. *Communication Education, 58,* 155–188.

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics and marriage.* New York: W. W. Norton & Company (Originally published 1985).

Escoffery, D. S. (2006). *How real is reality TV? Essays on representation and truth.* Jefferson, NC: McFarland & Company.

Frank, M. G., & Svetieva, E. (2013). Deception. In D. Matsumoto, M. G. Frank, & H. S. Hwang (Eds.), *Nonverbal communication: Science and applications* (pp. 114–132). Thousand Oaks, CA: Sage.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol 3, speech acts* (pp. 41–58). New York: Seminar Press.

Hall, A. (2006). Viewers' perceptions of reality programs. *Communication Quarterly, 54*(2), 191–211.

Hooked on Houses. (2012, June 11). *"House Hunters": What it was like to be on the show.* Retrieved from https://hookedonhouses.net/2012/06/11/house-hunters-what-it-was-like-to-be-on-the-show/.

Hubbell, A. (2014). Medical and pharmaceutical industry. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 651–653). Thousand Oaks, CA: Sage.

Kalbfleisch, P. J. (1990). Listening for deception: The effects of medium on accuracy of detection. In R. N. Bostrom (Ed.), *Listening behavior: Measurement and application* (pp. 155–176). New York: Guilford.

Kalbfleisch, P. J. (1992). Deceit, distrust and the social milieu: Application of deception research in a troubled world. *Journal of Applied Communication Research, 20,* 308–334.

Kalbfleisch, P. J. (1994). The language of detecting deceit. *Journal of Language and Social Psychology, 13*(4), 493–520.

Kalbfleisch, P. J. (2001). Deceptive message intent and relational quality. *Journal of Language and Social Psychology, 20*(1–2), 214–230.

Knapp, M. L. (2008). *Lying and deception in human interaction.* Boston, MA: Pearson Education.

Knapp, M. L., McGlone, M. S., Griffin, D. L., & Earnest, B. (2016). *Lying and deception in human interaction* (2nd ed.). Dubuque, IA: Kendall Hunt.

Konnikova, M. (2016). *The confidence game: Why we fall for it every time.* New York: Viking/Penguin.

Kurylo, A. (2014). Public relations. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 785–788). Thousand Oaks, CA: Sage.

Levine, T. R. (2014a). Truth-Default Theory (TDT): A theory of human deception detection. *Journal of Language and Social Psychology, 33*(4), 378–392.

Levine, T. R. (2014b). Frank, Mark. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 393–394). Thousand Oaks, CA: Sage.

Mazur, M. A., & Kalbfleisch, P. J. (2003). Lying and deception detection in television families. *Communication Research Reports, 20,* 200–207.

Pallegedara, R., & Hancock, J. (2014). Drug advertising. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 312–313). Thousand Oaks, CA: Sage.

Sandler, I. (2014). Vrij, Aldert. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 923–924). Thousand Oaks, CA: Sage.

Stearns, S. (2014a). Bok, Sissela. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 86–86). Thousand Oaks, CA: Sage.

Stearns, S. (2014b). Context. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 215–216). Thousand Oaks, CA: Sage.

Sternglanz, R., Morris, W., & Makiyil, J. (2014). DePaulo, Bella. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 285–286). Thousand Oaks, CA: Sage.

Strecker, E. (2012, June 12). 'House Hunters' scandal: Is the show a fake? *Entertainment Weekly.* Retrieved from http://ew.com/article/2012/06/12/house-hunters-fake-scandal-hgtv/.

Summary, J. (2014). Knapp, Mark. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 576–577). Thousand Oaks, CA: Sage.

Tsay-Vogel, M., & Krakowiak, K. M. (2017). Exploring viewers' responses to nine reality TV subgenres. *Psychology of Popular Media Culture, 6*(4), 348–360.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice.* Chichester: Wiley.

Wiking, M. (2016). *The little book of hygge: Danish secrets to happy living.* New York: HarperCollins.

Lie Catchers: Evolution and Development of Deception in Modern Times

Anne Solbu and Mark G. Frank

Research on lie detection shows that we are no better than chance at detecting lies (Bond & DePaulo, 2006). One reason may be that in day-to-day life we tend to believe that others are being truthful, which is referred to as *truth bias* (McCornack & Parks, 1986). There are different explanations regarding this tendency to take people at face value. Some researchers suggest people are naïve (e.g., O'Sullivan, 2003); others propose that people do not have enough information to make accurate judgments (Street, 2015); whereas others suggest people simply prefer to facilitate cooperation with others by assuming all utterances are true (Grice, 1975; Levine, 2014). However, in order to fully explain why people, in general, are not astute lie detectors, and to understand why there are exceptions to this tendency, we need to consider what we actually mean by lying and how it relates to our evolutionary background. We propose that widening the view of deception detection from the more proximal factors of a lie catcher's skill set, to those more distal factors involved in our development as a species, as people, and our social structure, will reveal that there was not much evolutionary pressure for us all to become good lie catchers individually, that some individuals are more apt at being good lie detectors, and that we as a collective may be sufficiently efficient at lie catching. We do this by examining phylogeny (our species), ontology

A. Solbu (✉) · M. G. Frank
Department of Communication, University at Buffalo,
State University of New York, Buffalo, NY, USA
e-mail: as255@buffalo.edu

M. G. Frank
e-mail: mfrank83@buffalo.edu

(our development), and sociology (our social structures). Specifically, we first compare humans and nonhuman primates to gain important insight into the development of the cognitive processes employed in decision-making (Santos & Rosati, 2015; Stevens, 2008), and deception (Bond & Robinson, 1988). Second, we examine children's development, to ascertain the cognitive steps required to detect lies and the importance of the social environment in the judgment process. Finally, we consider the elements of social evolution to better identify the role of the community in the detecting of lying.

PHYLOGENY

Defining Deception and Lying

The first step for better insight into the lie detection process is to detangle what we mean by lying and deception, because these concepts are often used interchangeably in a human only context (i.e., Vrij, 2007). When we examine deception phylogenetically, with a focus on the evolution in complexity of life forms from single cell to conscious beings, we see a corresponding evolution in complexity of deception, as well as a parallel evolution in complexity of detecting deception (as the ability to lie also implies ability to detect lies; Bussey & Grimbeek, 2000; Wright, Berry, & Bird, 2012). We argue that by understanding the range of sophistication in deceiving, from simple lower level appearance (i.e., 'looking like something else') to higher-order actions (i.e., movements and actions that are specifically intended to mislead based upon another's point of view; Mitchell, 1986), we will not only be better equipped at understanding deceiving and lying but also the lie detection process, including more closely identifying the sources of our lie detection biases.

Mitchell (1986) provides a phylogenetic model that illustrates this hierarchy of complexity. He proposes that there are four levels of deception found in nature, ranging from the simplest form to more advanced, higher-order lying. The *first* level consists of mimicking. Mitchell (1986) used the logical statement *always do x* (p. 29) as the deception strategy. By this, Mitchell (1986) means the deception is inherent in the appearance of the organism, thus it is always 'doing'—or better stated, 'living' its deception. For instance, the viceroy butterfly looks so similar to the monarch butterfly, that it fools the viceroy's predators into thinking it is the (apparently) less palatable monarch. Another example is when plants feature parts that look like a female bee's reproductive organs, thus attracting male bees for the purpose of pollination. At this level, the deception is in the appearance of the plant or animal and does not change for the duration of the plant or animal's life or in response to any specific actions on the part of other plants or animals. The only way it does change is in response to selective pressures over the course of many generations. For example, the peppered moth has white and black versions, and the white version predominated as it better deceived predators by blending into the light-colored bark of birch trees. However, as industrialization

generated pollution that darkened the bark of the birch trees, the black version predominated as that now better deceived predators (Cook, Grant, Saccheri, & Mallet, 2012). At this level of deception, according to Mitchell (1986), deception occurs from the perspective of the observer who is deceived at a higher level. For instance, a blue jay may learn to avoid the viceroy butterfly after experiencing nausea from eating the monarch. It may be very hard to detect this deception. There is no contemporaneous movement, or signal, for the lie catcher to observe outside of maybe minute structural differences. Typically, detection may only happen accidentally. For instance, although the blue jay has learned to avoid both butterflies, it may reflexively snatch the monarch butterfly by mistake, only to realize it was not a mistake and in fact was the quite palatable viceroy butterfly.

At the *second* level of deception, the organism has a reflex response to an external stimulus. Mitchell (1986) expanded the logical statement to read *do certain behavior x, when condition y is present* (p. 29) as the deception strategy. For instance, some snakes may feign death when disturbed (Mitchell, 1993). A rabbit may freeze, a possum may play dead, and a chameleon passively blends into the colors of its immediate environment (through chromatophores in the skin that absorb and reflect background light to change its color). In these instances, these reflexes are specific actions, but only in response to specific stimuli or challenges. We can presume that these reflex responses would be an evolutionarily derived action pattern, as distinct from the first level's evolutionary derived appearance. The first two levels are not all that relevant to humans, with rare occasions. For example, a child may do something bad, then go hide in a corner and cover his or her eyes (Mitchell, 1993). At this level of deception, there are behavioral signals that can be detected, and it would be up to the detector to recognize the differences between real death and fake death, or movement signs of the chameleon that gives away its camouflaged location, and so forth. No inferences about mental states of the organism are needed for the lie detector to be successful.

At the *third* level of deception, Mitchell (1986) describes behaviors that appear to be learned through an action that deceived another in the past. Mitchell's logical statement here is *do any x behavior given that x had resulted in y in your past* (p. 29). For example, a male green frog may learn to lower its pitch to give the illusion it is bigger, thus more formidable, than it actually is in order to defend its territory (Bee, Perrill, & Owen, 2000). A lion cub learns from its mother to observe prey motionless, and then approach slowly, crouched low in the grass, until it is close enough to risk breaking cover and running. It learns these actions deceive its prey, and as such the acts are intentional. However, in this instance we cannot infer anything about the mental state of the deceiver, and whether it has any thoughts whatsoever about the mental state of the deceived. Thus, the behaviors are intentional, but whether they are done with regard to intentionally changing the impressions in the minds of the other, or simply view the other as an object, is not always clear. Scholars who study animal deception carefully note that they are

examining behaviors that *appear* to be lying because they *look* intentional, but they cannot say it was definitely a lie, because they cannot determine if it was definitely deliberate (Premack, 2007; Towner, 2010; Whiten & Byrne, 1988). One cannot ask (or, at least get a verbal response from) a chimpanzee what its intentions were when it hid the bananas when out of sight of the other chimpanzees. The implications for detecting such deception are the similar, but not identical, to the second level, but with a bit more processing on the part of the detector. First, a detector may need to know the typical behavior of a given organism, not the species in general. The detector would be required to recognize that the counterfeit big green frog's song is not the same song as per usual from that specific green frog. Thus, a deception detector must not only look for subtle signs or signals that differentiate the real from the fake across species, but the change in behavior within a member of the species.

Only at the *fourth* level of deception does Mitchell (1986) propose that the sender intentionally attempts to manipulate the mental states of the receiver, such that the receiver accepts the false belief the sender is trying to instill. At this higher-order level of deception, the deceiver is capable of planning the act of deception through pretense and predicts its effect upon the thoughts or beliefs of the target of the lie (Mitchell, 1986, 1987). The logical statement from Mitchell (1986) is simply that each organism *self programs* its lie strategy (p. 29). Humans who lie are deceivers at this level. We note that most social scientists who have defined *lying* feature intentionality, or similar words such as *conscious, knowingly, deliberate,* as the key concept (Buller & Burgoon, 1996; DePaulo et al., 2003; Ekman, 1985/2001; Knapp & Comadena, 1979). For instance, Ekman (1985/2001) defined lying as "…one person intends to mislead another, doing so deliberately, without prior notification of this purpose, and without being explicitly asked to do so by the target" (p. 28). Similarly, there are many scholars who study animal and plant deception (i.e., Bond & Robinson, 1988; Mitchell, 1986; Trivers, 1985), but none suggest it is a conscious act, but simply something that serves a survival or reproductive function (Dawkins, 1976). To lie at this level requires the ability to envision how any given other sees the world, so the deceiver can act in a way to change the perceptions of that potential deception detector. This then shifts a heavier burden onto the lie catcher per se; he or she must try to understand the behavior seen, examine it not only in terms of the normal behavior of the individual, but also its relation to the context, as well as the motives of the person who is seemingly trying to deceive. The lie catcher must necessarily engage in higher-order cognitive processing to detect these lies.

Thus as these levels of deception become more cognitively and socially sophisticated, we note that the lie catchers also have to match the increased levels of cognitive and social sophistication. At the first two levels of

deception outlined by Mitchell (1986), simply detecting the organism would be enough to make accurate judgments. One would not need the ability to interpret. At the third level of deception, learning specific behavior is required in addition to the detection. However, at the fourth level, now the lie catcher cannot just detect the behavior (e.g., signs of nervousness in face and voice), but must interpret why this individual may be nervous, including honest reasons why (e.g., afraid of being disbelieved; Ekman, 1985/2001). Therefore, the area in which we should focus our attention on lie catching would seem to be these third and fourth levels of deception. We can glean this insight by examining primates.

Nonhuman and Human Primates

The ability of human beings to intentionally deceive is proposed to be related to advanced forms of cooperation, as higher level of social cognition is required for both (Baron-Cohen, 1999). Cooperation, or sharing plans and goals, does not merely imply a coincidental alignment of goals, but the recognition that another's goals are the same (or different) as one's own (Baron-Cohen, 1999). Advanced cooperation was proposed to emerge with the need to track dominance ranks and maintain complex social relationships (e.g., Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012); this in turn pushed for greater concern for others and perspective-taking (DeWaal, 2008). The social cognitive abilities associated with perspective-taking are known as Theory of Mind (ToM; Premack & Woodruff, 1978). ToM proposes that a being has an understanding that other beings have intentions, desires, and beliefs, and will act accordingly based upon the information before them (Wellman, 1992). As described earlier, a lie is designed to purposely alter another's knowledge to be false; in order to do this, a liar must have a ToM to know how the target of their lie will interpret the information presented before them.

Nonhuman primates, on the other hand, are proposed to have difficulty distinguishing their own mental states from others' (Tomasello, 1999). They may lack the ability to attribute false beliefs, as they are unable to understand that others can have a belief that does not correspond to reality (Suddendorf & Whiten, 2001). For instance, a monkey might not signal to other monkeys the presence of food in order to get more for themselves. However, monkeys rarely signal the presence of food when there is none, as this would require higher-order understanding of others' mental states (Premack, 2007). When food is present, and the monkey who sees it does not signal, but is caught with the food, that monkey is subject to being chased and is the target of acts of aggression by other group members (Hauser, 1992). The punishers act selfishly though (Jensen, 2010), evidenced by the ceasing of the chase by the attacker if food is dropped (Hauser, 1992). This suggests that the monkey's

actions during both deception and deception detection are driven by basic personal goals (e.g., to obtain more food) rather than considering the goals of the community.

Deception detection and related punishment among humans are more complicated on both a social level and a cognitive level due to the added complexity of human social life. The human ability to separate oneself from others, and to distinguish personal goals from the goals of society (e.g., Boyer, 2001), has allowed us to take others' perspectives to decide what punishment is required for the maintenance of social cohesion, despite not being personally affected by the lie or the liar (Gintis, 2000; Hall & Brosnan, 2017). This type of altruistic third-party punishment has not been evidenced in primates (Jensen, 2010).

Yet, great apes have some abilities related to ToM, specific to the simulation of others' actions (Suddendorf & Whiten, 2001). In an updated version of the levels of deception, Mitchell (1993) emphasizes pretense of others' actions and planning as a level preceding higher-order deception to better accommodate this intermediate step. The increased neocortical brain volume of humans compared to primates, and primates compared to other nonhuman animals, has been interpreted as evidence for higher level of social cognition across these species (Byrne & Whiten, 1988; Humphrey, 1976). Higher order of social cognition allows for the potential for larger group size (Dunbar, 1993), because more members require any given member to keep track of more social interactions. This also entails a better ability to respond with more nuanced social behaviors to reduce or avoid conflict created by the increase in members. Some researchers conclude that great apes (chimpanzees, orangutans, gorillas) have a limited ToM (Suddendorf & Whiten, 2001) and hence some precursor of higher-order deception (Courtland, 2015; Mitchell, 1993), likely comparable to a 2-year-old human (Suddendorf & Whiten, 2001). Regardless, it appears that the level of great ape ToM is such that they may be able to perpetrate deception, and deception that appears to be lying—deliberate, chosen, with some rudimentary understanding of what is going on in the minds of other great apes. However, their detection of such deception seems to be primarily based upon witnessing the reality, or 'catching the other in the act' (Hauser, 1992). Interestingly, human inclination to social learning, and overthinking, allows us to rationalize the behaviors of others, even those caught in the act, thus increases the risk of maintaining inaccurate social opinions for longer duration, compared to other species (Rauwolf, Mitchell, & Bryson, 2015; Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009).

Although many differences, there is thus evidence of some, albeit minor, overlap between human and great apes' ability to deceive and thereby mental ability to detect deception. By taking into account and comparing a more continuous set of psychological mechanisms in human and nonhuman primates (Towner, 2010), it is possible to better pinpoint the origins of human decision-making and biases (Santos & Rosati, 2015).

Ontology

Children and Adolescents

Understanding the development of cognitive abilities is crucial in the study of evolution and decision-making (Stevens, 2008). Talwar and Lee (2008) outline a model of children's lying including three stages of development. During the *primary* stage, occurring between 2 and 3 years of age, children begin to occasionally produce false statements to avoid punishment, protect themselves, or be regarded more positively (Newton, Reddy, & Bull, 2000; Wilson, Smith, & Ross, 2003). These lies are considered rudimentary, as the intent of instilling a false belief into others is uncertain (Talwar & Lee, 2008). Children in this stage can be considered to have acquired an 'early' ToM (Chandler, Fritz, & Hala, 1989) as they begin to understand that others can have diverse desires and beliefs, and that others may not know what they know, or what is true (Leduc, Williams, Gomez-Garibello, & Talwar, 2017; Ma, Evans, Liu, Luo, & Xu, 2015; Wellman & Liu, 2004). One study found that a child's ability to comprehend the concept of knowledge access, along with their ability to inhibit themselves from telling the truth, predicted lie-telling in children between 2 and 4 years of age (Leduc et al., 2017). In general, the most basic types of lying in children involve simple denial, as opposed to rich fabrication, and are self-serving in nature, likely requiring less mature ToM than later occurring pro-social lying (Williams, Moore, Crossman, & Talwar, 2016).

Children enter the *secondary* stage of lying development around the age of 4, when the majority begin to lie to cover up transgressions in more refined ways, indicating that they are able to predict what someone else is feeling or thinking (Talwar & Lee, 2008), also known as having *first-order belief understanding* (Wimmer & Perner, 1983). Relatedly, children begin to attribute false beliefs, as they understand that others can have a belief that does not correspond to actuality (Astington, 1993; Wellman & Bartsch, 1988). Children at this stage become more successful in appearing honest in their nonverbal behavior, yet have difficulty maintaining their lies during questioning, also referred to as *semantic leakage* (Talwar & Lee, 2002). It is not until the age of 7 or 8 when children enter the *tertiary* stage of lying development that they are able to control for this semantic leakage. This is facilitated by the acquisition of second-order belief understanding (Talwar & Lee, 2008), which is the ability to predict what someone thinks or feels about someone else's thoughts or feelings (Perner & Wimmer, 1985).

As children's skill in lying emerges, it seems their ability to detect lies and truths develops in parallel (Bussey & Grimbeek, 2000). Children under the age of five have difficulties identifying lies but are still able to correct false statements, indicating that they have the cognitive abilities required for lie detection but fail because they simply expect others will be honest (Mascaro, Morin, & Sperber, 2017). Children have also shown difficulty identifying

harmful truths (Talwar, Williams, Renaud, Arruda, & Saykaly, 2016). This may stem from strong parental dependency (Mascaro et al., 2017), as well as early socialization to believe that lying is immoral and sinful (Piaget, 1932/1965). In general, the belief that all forms of lying are wrong lasts until around age 10, when a child's sensitivity to a liar's intention increases enough that they can factor that into their judgment (Piaget, 1932/1965); although some have shown it can even be a few years earlier than 10 (Xu, Bao, Fu, Talwar, & Lee, 2010; Xu, Luo, Fu, & Lee, 2009). At this point, children develop a more favorable evaluation of pro-social lies (lies told to protect others from harm; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996) compared to selfish lies (i.e., Bussey, 1999; Talwar et al., 2016). This appears to co-occur with the acquisition of second-order belief understanding (Cheung, Siu, & Chen, 2015). By the time children become adolescents, they will rate false statements with pro-social motives as less negative than statements with malevolent motives (Lee & Ross, 1997). This may be due to a more developed ToM caused by peer influence, which allows them to integrate multiple mental states from many individuals (Kuhn, 2000).

Altogether, the social intention of the communicator, in addition to the maturity of the child, both play a role in the evaluation of statement veracity and magnitude of reward or punishment (Talwar et al., 2016). However, although children follow similar developmental stages as outlined in Mitchell's (1986) levels of deception, there are individual differences. Some children seem very interpersonally perceptive, others oblivious, and some on the autism spectrum virtually immune to nuances of behavioral interaction in general (Baron-Cohen, 1999). These differences typically translate into adulthood.

Individual Differences in Lie Detection

The role of the development of the mind on lie detection may not only be evidenced by distinguishing between humans and primates, and in the growth of children's ToM, but also between individuals.

Regarding lie detection in specific, scholars have identified superior ability in detecting lies in some individuals (Bond, 2008; O'Sullivan & Ekman, 2004) and professional groups, such as Secret Service professionals (Ekman & O'Sullivan, 1991; Ekman, O'Sullivan, & Frank, 1999), clinical psychologists with higher level of motivation to learn about lie detection (Ekman et al., 1999; O'Sullivan & Ekman, 2004), and even law enforcement in general, if they are presented with more ecologically valid material (O'Sullivan, Frank, Hurley, & Tiwana, 2009). One factor that seems to make individuals better lie catchers is their ability to maintain an open mind so that they do not apply hard and fast rules to evaluate every person (Ekman & O'Sullivan, 1991). Research suggests that when areas of the brain important to ToM are stimulated via transcranial currents—areas presumed to be involved in perspective-taking—it improves lie detection accuracy when faced with opinions

that conflicted with the participant's own opinions (Sowden, Wright, Banissy, Catmur, & Bird, 2015). Susceptibility to emotional changes in others at an early age may also improve lie detection ability (O'Sullivan & Ekman, 2004). Accurate lie detectors report relying on more nonverbal behavior (Bond, 2008; Ekman & O'Sullivan, 1991), or a combination of verbal and nonverbal behavior, but not on verbal behavior alone (Ekman & O'Sullivan, 1991). Training to improve detecting deception seems possible even if the methods are sub-par (Frank & Feeley, 2003). Others found training on identifying expression of emotion can increase lie detection success (e.g., Shaw, Porter, & ten Brinke, 2013). The importance of facial expressions in lie detection is further supported by the finding that aphasics, people with left hemisphere brain lesions preventing them from comprehending speech, are significantly more accurate than control groups at detecting lies (Etcoff, Ekman, Magee, & Frank, 2000). In contrast, persons with poor nonverbal processing, such as those with autism, have a reduced ability to detect (and produce) lies (e.g., Sodian & Frith, 1992).

Although there are links between the development of lying and lie detection (Bussey & Grimbeek, 2000; Wright et al., 2012), additional mechanisms, such as adherence to pro-social norms, or empathy appear to come into play in the practical usage of these abilities. For instance, one study found that socially skilled individuals were better able to deceive than socially anxious individuals (Riggio, Tucker, & Throckmorton, 1987), but another study showed that high social intelligence prevented accurate deception detection due to increased compassion (Baker, ten Brinke, & Porter, 2013). Conversely male, but not female, psychopaths with interpersonal exploitative tendencies where they tell many lies were found to be accurate at detecting deception (Lyons, Healy, & Bruno, 2013). Relatedly, depressed individuals have a tendency to view the world more accurately (Alloy & Abramson, 1979), whereas inaccurate views may increase happiness (Cummins & Nistico, 2002).

Therefore, it appears that a few developmental principles are predictive of good lie catchers at an individual level. Those with better ToM abilities, those with open minds, and those with better nonverbal detection and assessment abilities seem to outperform their peers. It is an interesting question as to why they are so good—is it that these individuals are significantly better than others due to some super skill set they possess, or is it the case that everyone else, who are barely above chance at accurately detecting lies, are being suppressed by social or cultural factors?

Sociology

Sociocultural Context

Social and cultural norms govern much of our interpersonal behavior. One of the most important sociocultural norms is that we should not state what

we believe to be false in conversations (Grice, 1989). We find this norm for honesty stretches across the world and cultures (Knapp, 2008). We are socialized into endorsing this norm, through cooperation and politeness, from childhood (Saarni & Weber, 1999). Nonetheless, culture-specific conventions regarding lying are very important when assessing how we evaluate honest and dishonest interactions (Dor, 2017; Lee, 2000; Sweetser, 1987). This is because we interpret messages based not only on the perceived intention of the message and the messenger, but also on the sociocultural context in which it occurs (Lee, 2013). For instance, a message may be classified into four categories: as cooperative honesty, harmful honesty, cooperative lying, or harmful lying (Dor, 2017). However, to determine the specific intention (i.e., to harm or to cooperate), we must understand the cultural context (environment). For example, in Japan, lying is often utilized as a means to suppress oneself and adhere to the rules of society (Freeman, 2009). Thus when Japanese and American participants were exposed to horrific films, the Japanese were more likely to smile and mask their negative emotions in the presence of a high-status person (Friesen, 1972). This was done to adhere to a Japanese norm of not showing negative emotions to a high-status person, thus covering these emotions with a smile was done to be seen as polite (Friesen, 1972). If one was unfamiliar with the cultural context, he or she may misinterpret this as un-cooperative lying, when in fact it was cooperative lying.

Similarly, white lies are deliberate misinformation usually uttered for polite reasons, often to save someone from harm or to preserve their feelings; for example, the white lie a guest utters when he or she tells the host the meal was great when in fact the guest did not like the food (Sweester, 1987). Yet one study cautioned that in some cultures, such as that found in Ecuador, lies were not only generally rated as less acceptable by Ecuadorans than Americans of European decent, but even white lies were viewed negatively (Mealy, Stephan, & Urrutia, 2007). Subcultures may also view the acceptability of lies differently; for example, members of the Church of Jesus Christ of Latter-Day Saints rated lies as less acceptable than non-members (Ning & Crossman, 2007). The cultural view of what counts as lying may be so strict that even the deliberateness of the attempt to mislead subsumed in the message no longer matters. For example, the Mopan Maya of Southern Belize view all untrue statements as lying, regardless of if they are intended to mislead or not (Danziger, 2010).

Generally, lying to enemies or the out-group is more accepted than lying within the group (Dunbar et al., 2016; Mealy et al., 2007; Sweester, 1987). This effect appears to be especially strong in collectivist cultures (see Fu, Evans, Wang, & Lee, 2008). For example, researchers found that in China, children supported *blue lies* (i.e., lies told on behalf of a group, which often endears the liar to the group) but viewed truths against their group unfavorably (Fu et al., 2008; Lee, 2013). Interestingly, these sociocultural views on lying do not only translate to evaluation of lying but also to production

of lying. The children's endorsement of blue lies was positively related to their actual lying behavior (Fu et al., 2008) and it has also been found that Samoans, but not Americans, attempt to lie for collectivist purposes (Aune & Waters, 1994).

In some cultures, such as that found in Lebanon, even if lying in general is seen as immoral, successful lying by men can be seen as status enhancing, while at the same time prestige is often given to successful male lie detectors (Gilsenan, 1976). This may be due to the prominence of masculine dominance found in that culture, where power and material success of men are important (Hofstede, 1980). In other situations, the way lies are perceived may be a function of power status in a culture or group between the powerful and their subordinates (see Hofstede, 1980). In one study, four- and five-year-old Catholic Italian children would never believe that a priest told a lie (Fu et al., 2008).

Taken together, the culture and social context transcend into our expression and evaluation of lying, which in turn can affect the ability to detect a lie. We recognize that, language barriers aside, cross-cultural lie detection is more difficult than intra-cultural lie detection. This was evidenced in a cross-cultural study with both American and Jordanian liars/truth tellers and lie detectors, where between-culture judgment accuracy was below 50%, despite within-culture accuracy of slightly above chance (Bond, Omar, Mahmoud, & Bonser, 1990). Some cultural behavioral styles lend themselves to judgments of deception whether or not the individual is lying; for example, Afro-Dutch in Suriname are less likely to make eye contact, causing them to be more likely to be judged as lying by the Dutch (Vrij, Dragt, & Koppelaar, 1992). Other cultural norms regarding postures, emotion expression regulation, and so forth can also lead to misjudgments of truth or lie (Efron, 1941; Ekman & Friesen, 1969; Hall, 1966). In general, we could expect more honest communication within groups (Fitch, 2010) and more lying between groups (Knight, 1998), reflecting veracity judgments similarly. Finally, what is considered a white lie may vary from culture to culture. White lies are a social lubricant to ease social interaction and therefore are not to be scrutinized by the lie catcher (Frank & Svetieva, 2013). The lies we want to catch are the harmful lying types; yet if we are uncertain as to the social norms, and what constitutes harmful lying, versus cooperative lying, we can see how we may be unprepared to examine almost all utterances for veracity. This can only work against developing any particular skills toward lie detection.

Social Evolution

Humans evolved the ability to engage in symbolic thought—using sounds (language) to represent ideas or objects (McCrone, 1991). It is thus a short step to represent ideas or objects that are not present—that is, a proposition which is untrue—which is the foundation of a lie. Given the basic assumption

that communication is honest (Grice, 1975), it is possible for individuals to 'cheat' by stating this false information to obtain some advantage over others, also known as free riding (e.g., Gintis, 2000; Krebs & Dawkins, 1984; Trivers, 1971). What this means is that within the overarching environment of cooperative, honest communication there are opportunities for tactical deception (McNally & Jackson, 2013). In essence, tactical deception involves utilizing normal behavior in a different context, such that the act is misinterpreted, providing the deceiver an advantage (Byrne & Whiten, 1991). It is further believed that not only did language and the evolution of the mind contribute to deception, but deception contributed to the evolution of the mind (e.g., Byrne & Whiten, 1992) and language (Dor, 2017). For example, successful liars, and also good lie detectors, would have been able to gain a competitive advantage and thus produce more offspring (e.g., Bond & Robinson, 1988; Dawkins & Krebs, 1979).

Tactical deception fits into the context of the evolution of human cooperation and interdependence, also known as the *interdependence hypothesis* (Tomasello et al., 2012). This model proposes two steps in developing cooperation; in a first step, people developed joint intentionality, formed collaborative partnerships, identified roles, and generally helped each other and avoided potential cheaters. In the second step, as the modern human population grew, groups competed for resources and in the process further developed cognitive skills serving the collective, also labeled *collective intentionality*, leading to the formation of cultural norms and practices (Tomasello et al., 2012).

Tactical deception or 'free riding' was likely not a frequent occurrence, hence it was not a rampant social problem, in the first step because collaboration of everyone in the small group was necessary for success. Collaboration then implied retribution or punishment for those who violated the collaborative norms, which in turn encouraged cooperation (Tomasello et al., 2012). Thus trusting others, and taking each other at face value, was fundamental to the success of the group, even if it entailed some minor exploitation by free riders (Dunbar, 2004).

Game theory approaches also confirm that small amounts of tactical deception would not disrupt the social system. Game theory in general is based upon the assumption that there are differential outcomes to individuals who cooperate (altruistic, truthful) or compete (selfish, deceive) with each other during social interactions; one form of game theory is the prisoner's dilemma (Tucker, 1950). In the prisoner's dilemma, if both individuals choose to cooperate, they both benefit; if they both chose to compete, they both suffer, if one competes, and the other cooperates, then the cooperator suffers greatly. When iterative (repeated) interactions are computer simulated, the results of those studies initially suggest that over thousands of such interactions, the person who chooses to compete succeeds more than the one who chooses to cooperate, and ultimately eliminates the cooperator. Thus, it pays off better to be a competitor compared to a cooperator. However,

this finding seemed to be at odds with our generally cooperative social life. Scholars rectified this apparent discrepancy by re-running those iterative models, but this time building in the concept of sanction (i.e., discovering the person was a competitor, and then refusing to engage the competitor). When sanction was entered, the models now revealed that those who choose to cooperate, who then sanction known competitors, would succeed more. Thus, it now pays off better to be a cooperator compared to a competitor (as they would have eliminated the known competitors, leaving only the cooperators to play; Cosmides & Tooby, 1989). Therefore, small amounts of cheating may go undetected; large amounts will eventually be detected and thus the competitor would be ostracized. The first evolutionary step then suggests we did not need super-perceptive abilities to catch every lie, but a person who lied too much for their own benefit would eventually get caught, and the consequences in the ancient world of being ostracized would be severe or even fatal.

Free riding became a bigger problem in the second step. As group sizes increased and society spread, it was more difficult to obtain information regarding reputation (Dunbar, 2004; Enquist & Leimar, 1993; Tomasello et al., 2012). This led to an increased necessity to rely on memory regarding free riders' past transgressions, ability to predict infractions affecting others, and ultimately language to communicate these concerns (Dunbar, 2004). Scholars have referred to the process in which information regarding people's reputations is exchanged as gossip (e.g., Enquist & Leimar, 1993). Gossip is thought to promote pro-social behavior both by excluding selfish individuals and by deterring un-cooperative acts (Willer, Feinberg, Irwin, Schultz, & Simpson, 2010), as individuals do not want to stand out (Sommerfeld, Krambeck, & Milinski, 2008). Similarly, impression management (Engelmann, Herrmann, & Tomasello, 2012) or 'relatedness' motives—to make one's self appear to be more similar to the group—became more important to encourage smooth interactions (Haidt, 2001).

Altruistic punishment for detected free riders, which benefit the larger group, likely did not develop until societies became large and anonymous, facilitating so-called *hyper-cooperation*, where all behavioral activities were seen and understood as beneficial to the group (Burkart et al., 2014). Without the implementation of punishment, cooperation would have dropped as anonymity grew (e.g., Franzen & Pointner, 2012; Haley & Fessler, 2005; Hoffman, McCabe, Shachat, & Smith, 1994). Norms were more easily internalized and cooperation became 'instinctive' when punishment of free riders was implemented, executed by a minority of enforcers (Gavrilets & Richerson, 2017). This is consistent with the social intuitionist model, which suggests that our moral judgments are effortless and in part intuitive, promoting social harmony at the expense of accuracy (Haidt, 2001).

Taken together, the evolution of human social structures, including collaboration and group interdependence, produced a backdrop where members

could safely assume that communication was honest most of the time. Continual assessment of every utterance was not needed as close interaction, and gossip, served a surveillance function to keep tabs on each member to insure they were in good standing as a cooperator. Selfish lies (vs. white lies) were rare, as the cost of discovery was high, and the individual who told too many would eventually be unmasked by the group network.

Societal Fluctuations and Lie Detection in Modern Times

The societal 'radar' system of surveillance is not foolproof at any given point in time. As the social structures harden, liars have to become more sophisticated, and in return the lie catchers also have to become more sophisticated to keep this balance. The continuous adaptation in order to survive coevolving opposing forces in a dynamic environment, also known as the Red Queen effect (Van Valen, 1973), may explain the continuous coevolving spiral of deception and deception detection (Rauwolf, 2016). It is like an arms race, where each new measure must be matched by a new countermeasure (Dawkins & Krebs, 1979). Social interaction would then seem to coevolve cooperation and small amounts of selfish manipulation, but only so much that most selfish behavior can remain undetected (McNally & Jackson, 2013).

In essence, strong social bonds enhance conformity, leading to cooperation benefits, with reduced necessity to understand and act on the truth regarding individuals' strategic motives as this may fragment the group (Rauwolf et al., 2015). This acceptance of truth is very energy efficient, as the cost of 24/7 vigilance for every statement or action is exhaustion (Ekman, 1996). This translates into interpersonal lie detection as well, for instance, it may explain why spouses (subconsciously) choose to turn a blind eye to the opposing spouse's cheating. In fact, the act of not knowing, or self-deception, may be beneficial in deceiving others in order to gain social benefits, for example, having others admire the apparent devoted couple despite the cheating spouse (Chance & Norton, 2015; Trivers, 1991; Von Hippel & Trivers, 2011). Similarly, only seeking out information in order to maintain current beliefs, also referred to as *confirmation bias* (Jonas, Schulz-Hardt, Frey, & Thelen, 2001; Nickerson, 1998; Schulz-Hardt, Frey, Lüthgens, & Moscovici, 2000), may reduce the necessity to deceive altogether, minimizing the risk of costly detection (Rauwolf et al., 2015).

However as cooperation rises so does trust in a society. As people trust each other more, the lack of scrutiny for cheating increases, leading to more prevalent free riding (e.g., Dawkins & Krebs, 1979; Feinberg, Willer, Stellar, & Keltner, 2012). Although the number of free riders will increase, in any society there will always be a smaller proportion of persons with blatant free riding which demonstrates a complete disregard for societal norms. In fact, although it has been reported that the average person tells 1–2 lies per day (DePaulo et al., 1996), another study found that the average of 1–2 lies

per day were produced by a few prolific liars, whereas the majority told none (Serota & Levine, 2015). Similarly, there are always a smaller number of individuals with an extremely strong sense of morality and community affiliation, who will continue to make sacrifices for the public good (Gavrilets & Richerson, 2017). These individuals could serve as the countervailing force to balance out the prolific liars, by being the prolific lie monitors/detectors. In a society with a large proportion of selfish individuals, pro-social morality monitors would redress the balance, thus again removing the need for all individuals to develop super precise detection ability.

The dynamics of manipulation versus detection and enforcement may be employed to explain modern events such as the #MeToo movement which is based upon coalition building, facilitated by reputational communication, or pro-social gossip (gossip that 'leaks' accurate information to the benefit of the larger group, e.g., government corruption), and increased level of empathy. During the course of evolution, females likely created alliances designed to guarantee mating with cooperative male hunters, and the exclusion or punishment of un-cooperative males, in order to ensure survival of the offspring (Knight, 1991, 2008). This phenomenon, where a coalition is put in place to oppose individual dominance, has been labeled *reverse dominance* (Boehm, 1993), and brought about the evolution of human trust (Sztompka, 1999) and norms (Knight, 2008). The #MeToo movement illustrates the power of collective lie detection by showing how an initial claim of 'cheating' or deception (or sexual assault) made at a given individual may serve to draw out further claims against that same individual from other aggrieved people, thus validating the initial claim and exposing this 'cheater' to the wider community. This communal lie detection is a concept lacking in recent theories formulated to explain human deception detection (Levine, 2014; Street, 2015). Moreover, movements such as #MeToo also show that collective gossip may deter future free riding or selfish cheating (see Wilson, Wilczynski, Wells, & Weiser, 2000), as substantiated allegations are often met with guilty verdicts not only in the court of public opinion but in the actual courtroom.

Thus, we see that deception and deception detection continue to coevolve. The emergence of technology and social media has brought about the ability to connect with and monitor a large amount of individuals, providing opportunities for deceivers across the world, but also for individuals to 'catch' a cheater (liar), to gossip, and bring out the cheater worldwide for sanction, as illustrated by the #MeToo movement and its use of social media.

Conclusion

By examining the evolutionary history and relationships among species (phylogeny) and among humans (ontology), in addition to our social structures (sociology), with a focus on deception detection, we discover parallel steps of increasing socio-cognitive development and Theory of Mind. These

abilities do not just enable humans to intentionally create false beliefs in others but also the ability to detect manipulation or deception (Sip, Roepstorff, McGregor, & Frith, 2008; Spence et al., 2004; Wright et al., 2012). Understanding ToM development then appears to be central for understanding lie detection; why we're often inaccurate, our biases, and how we make our decisions about lying in general.

Better insight into ToM may explain why excellent lie detection by some individuals is possible. Certainly, some groups or individuals can show high accuracy (O'Sullivan et al., 2009). Moreover, specific life situations, perhaps coinciding with early ToM development in childhood, can push for such abilities within a given individual; for example, someone raised in a violence fraught area, or someone being raised by an alcoholic, where they had to detect true intentions quickly, at great consequence to the individual's safety (O'Sullivan & Ekman, 2004). Thus, some people can be very good at detecting lies. And individuals can be trained to be better lie detectors (Frank & Feeley, 2003).

Nonetheless, studying the evolution of deception and lie detection, we also discover that a strong press on humans in general to be excellent lie detectors does not seem to be present. The ToM, our cooperative nature, along with our ability to rationalize based on other internal or social needs, means we can often refuse to properly interpret evidence of lying before our own eyes. It also means we can selectively lie but not too much; just enough to 'stay below the radar.' Thus, it appears that we are not pushed evolutionarily to develop a corresponding detection mechanism to catch every lie. Therefore, it makes sense as to why we get the repeated finding that individuals in general are not very good at spotting deception.

However, lie detection in the laboratory or in single case studies does not fully translate to the real world where gossip and relationships with others matter (Haidt, 2001). People rely on gossip, whether accuracy is limited (Sommerfeld et al., 2008), or actually improves lie detection (Klein & Epley, 2015). Moreover, it is through the influence from others that we may decide to override our tendency to cooperate (Bear & Rand, 2016), and employ conscious deliberation to make our decisions (Haidt, 2001). The alignment of emotions through empathy, and increased goal sharing (Tomasello, Carpenter, Call, Behne, & Moll, 2005), as evidenced by the #MeToo movement (Rodino-Colocino, 2018), gave rise to the same powerful group thinking and sociality as seen in the emergence of human morality (Jensen, Vaish, & Schmidt, 2014). Haidt (2001) states "A group of judges independently seeking truth is unlikely to reach an effective consensus, but a group of people linked together in a large web of mutual influence may eventually settle into a stable configuration" (p. 826). This becomes, functionally, a long-range radar type system that has agents reporting back actions, behaviors, relationships to each other, which in turn sets the groundwork for recognizing inconsistencies regarding people not being where they say they are, people being with people they deny knowing, and so forth. The presence of this communication

network would reduce the need to make individuals hyper-vigilant, or to individually develop super-acute deception detection skills. Likewise, unusual interpersonal behaviors can trigger individuals to search for evidence to verify their hypotheses about someone's veracity, and they can then activate their social networks to verify the information provided by the unusually behaving person (Novotny et al., 2018). Thus, these networks are not just passive providers of information. We believe the research literature has neglected this larger system in which our social structures exist, that often detect the deception for us. Even as our society expands, social media and movements like #MeToo have become like the global village, where previously unacquainted individuals can now verify the truth or falsity of each other, thus (hopefully) betraying the attempted liar.

References

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108,* 441–485. https://doi.org/10.1037/0096-3445.108.4.441.

Astington, J. W. (1993). *The child's discovery of the mind.* Cambridge, MA: Harvard University Press.

Aune, R. K., & Waters, L. L. (1994). Cultural differences in deception: Motivations to deceive in Samoans and North Americans. *International Journal of Intercultural Relations, 18,* 159–172. https://doi.org/10.1016/0147-1767(94)90026-4.

Baker, A., ten Brinke, L., & Porter, S. (2013). Will get fooled again: Emotionally intelligent people are easily duped by high-stakes deceivers. *Legal and Criminological Psychology, 18,* 300–313. https://doi.org/10.1111/j.2044-8333.2012.02054.x.

Baron-Cohen, S. (1999). Evolution of a theory of mind? In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 261–277). New York, NY: Oxford University Press.

Bear, A., & Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences, 113,* 936–941. https://doi.org/10.1073/pnas.1517780113.

Bee, M. A., Perrill, S. A., & Owen, P. C. (2000). Male green frogs lower the pitch of acoustic signals in defense of territories: A possible dishonest signal of size? *Behavioral Ecology, 11,* 169–177. https://doi.org/10.1093/beheco/11.2.169.

Boehm, C. (1993). Egalitarian behavior and reverse dominance hierarchy. *Current Anthropology, 34*(3), 227–254.

Bond, G. D. (2008). Deception detection expertise. *Law and Human Behavior, 32,* 339–351. https://doi.org/10.1007/s10979-007-9110-z.

Bond, Jr., C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14,* 189–204. https://doi.org/10.1007/BF00996226.

Bond, C. F., & Robinson, M. (1988). The evolution of deception. *Journal of Nonverbal Behavior, 12,* 295–307. https://doi.org/10.1007/BF00987597.

Boyer, P. (2001). *Religion explained: The evolutionary foundation of religious belief.* New York, NY: Basic Books.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x.

Burkart, J. M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., …, van Schaik, C. P. (2014). The evolutionary origin of human hyper-cooperation. *Nature Communications, 5,* 4747. https://doi.org/10.1038/ncomms5747.

Bussey, K. (1999). Children's categorization and evaluation of different types of lies and truths. *Child Development, 70,* 1338–1347. https://doi.org/10.1111/1467-8624.00098.

Bussey, K., & Grimbeek, E. J. (2000). Children's conceptions of lying and truth-telling: Implications for child witnesses. *Legal and Criminological Psychology, 5,* 187–199. https://doi.org/10.1348/135532500168083.

Byrne, R. W., & Whiten, A. (1988). Toward the next generation in data quality: A new survey of primate tactical deception. *Behavioral and Brain Sciences, 11,* 267–273. https://doi.org/10.1017/S0140525X00049955.

Byrne, R. W., & Whiten, A. (1991). Computation and mindreading in primate tactical deception. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 127–141). Cambridge, MA: Basil Blackwell.

Byrne, R. W., & Whiten, A. (1992). Cognitive evolution in primates: Evidence from tactical deception. *Man, 27,* 609–627. https://doi.org/10.2307/2803931.

Chance, Z., & Norton, M. I. (2015). The what and why of self-deception. *Current Opinion in Psychology, 6,* 104–107. https://doi.org/10.1016/j.copsyc.2015.07.008.

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development, 60,* 1263–1277. https://doi.org/10.2307/1130919.

Cheung, H., Siu, T. S. C., & Chen, L. (2015). The roles of liar intention, lie content, and theory of mind in children's evaluation of lies. *Journal of Experimental Child Psychology, 132,* 1–13. https://doi.org/10.1016/j.jecp.2014.12.002.

Cook, L. M., Grant, B. S., Saccheri, I. J., & Mallet, J. (2012). Selective bird predation on the peppered moth: The last experiment of Michael Majerus. *Biology Letters, 8,* 609–612. https://doi.org/10.1098/rsbl.2011.1136.

Cosmides, L., & Tooby, J. (1989). Evolutionary psychology and the generation of culture, part II: Case study: A computational theory of social exchange. *Ethology and Sociobiology, 10,* 51–97. https://doi.org/10.1016/0162-3095(89)90013-7.

Courtland, S. D. (2015). Detecting animal deception. *The Journal of Mind and Behavior, 36*(3/4), 121–138.

Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive cognitive bias. *Journal of Happiness Studies, 3,* 37–69. https://doi.org/10.1023/A:1015678915305.

Danziger, E. (2010). On trying and lying: Cultural configurations of Grice's Maxim of Quality. *Intercultural Pragmatics, 7,* 199–219. https://doi.org/10.1515/iprg.2010.010.

Dawkins, R. (1976). *The selfish gene*. New York, NY: Oxford University Press.

Dawkins, R., & Krebs, J. R. (1979). Arms races between and within species. *Proceedings of the Royal Society B, 205,* 489–511. https://doi.org/10.1098/rspb.1979.0081.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995. https://doi.org/10.1037/0022-3514.70.5.979.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118.

De Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59,* 279–300. https://doi.org/10.1146/annurev.psych.59.103006.093625.

Dor, D. (2017). The role of the lie in the evolution of human language. *Language Sciences, 63,* 44–59. https://doi.org/10.1016/j.langsci.2017.01.001.

Dunbar, R. I. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences, 16,* 681–694. https://doi.org/10.1017/S0140525X00032325.

Dunbar, R. I. (2004). Gossip in evolutionary perspective. *Review of General Psychology, 8,* 100–110. https://doi.org/10.1037/1089-2680.8.2.100.

Dunbar, N. E., Gangi, K., Coveleski, S., Adams, A., Bernhold, Q., & Giles, H. (2016). When is it acceptable to lie?: Interpersonal and intergroup perspectives on deception. *Communication Studies, 67,* 129–146. https://doi.org/10.1080/10510974.2016.1146911.

Efron, D. (1941). *Gesture and environment*. New York, NY: Kings Crown Press.

Ekman, P. (1985/2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. New York, NY: Norton.

Ekman, P. (1996). Why don't we catch liars? *Social Research, 63*(3), 801–817.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1,* 49–98. https://doi.org/10.1515/semi.1969.1.1.49.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46,* 913–920. https://doi.org/10.1037/0003-066X.46.9.913.

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10,* 263–266. https://doi.org/10.1111/1467-9280.00147.

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS One, 7*. https://doi.org/10.1371/journal.pone.0048433.

Enquist, M., & Leimar, O. (1993). The evolution of cooperation in mobile organisms. *Animal Behaviour, 45,* 747–757. https://doi.org/10.1006/anbe.1993.1089.

Etcoff, N. L., Ekman, P., Magee, J. J., & Frank, M. G. (2000). Lie detection and language comprehension. *Nature, 405,* 139. https://doi.org/10.1038/35012129.

Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: Reputational information sharing as prosocial behavior. *Journal of Personality and Social Psychology, 102,* 1015–1030. https://doi.org/10.1037/a0026650.

Fitch, W. T. (2010). *The evolution of language*. New York, NY: Cambridge University Press.

Frank, M. G., & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research, 31,* 58–75. https://doi.org/10.1080/00909880305377.

Frank, M. G., & Svetieva, E. (2013). The role of nonverbal behavior in detecting and telling lies. In M. L. Knapp & J. A. Hall (Eds.), *Nonverbal communication* (pp. 471–511). New York, NY: Mouton de Gruyter.

Franzen, A., & Pointner, S. (2012). Anonymity in the dictator game revisited. *Journal of Economic Behavior & Organization, 81,* 74–81. https://doi.org/10.1016/j.jebo.2011.09.005.

Freeman, D. M. A. (2009). Sociocultural perspectives on dishonesty and lying. In S. Akhtar & H. Parens (Eds.), *Lying, cheating, and carrying on: Developmental, clinical, and sociocultural aspects of dishonesty and deceit* (pp. 111–127). New York, NY: Jason Aronson.

Friesen, W. V. (1972). *Cultural differences in facial expressions in a social situation: An experimental test of the concept of display rules* (Unpublished doctoral dissertation). University of California, San Francisco.

Fu, G., Evans, A. D., Wang, L., & Lee, K. (2008). Lying in the name of the collective good: A developmental study. *Developmental Science, 11,* 495–503. https://doi.org/10.1111/j.1467-7687.2008.00695.x.

Gavrilets, S., & Richerson, P. J. (2017). Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences, 114,* 6068–6073. https://doi.org/10.1073/pnas.1703857114.

Gilsenan, M. (1976). Lying, honor, and contradiction. In B. Kapferer (Ed.), *Transaction and meaning: Directions in the anthropology of exchange and symbolic behavior* (ASA Essays in Social Anthropology Series, Vol. 1) (pp. 191–219). Philadelphia, PA: Institute for the Study of Human Issues.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206,* 169–179. https://doi.org/10.1006/jtbi.2000.2111.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics. Volume 3: Speech acts* (pp. 225–242). New York, NY: Seminar Press.

Grice, H. P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108,* 814–834. https://doi.org/10.1037/0033-295X.108.4.814.

Haley, K. J., & Fessler, D. M. (2005). Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26,* 245–256. https://doi.org/10.1016/j.evolhumbehav.2005.01.002.

Hall, E. T. (1966). *The hidden dimension.* New York, NY: Doubleday.

Hall, K., & Brosnan, S. F. (2017). Cooperation and deception in primates. *Infant Behavior and Development, 48,* 38–44. https://doi.org/10.1016/j.infbeh.2016.11.007.

Hauser, M. D. (1992). Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: A learned mode of production? *The Journal of the Acoustical Society of America, 91,* 2175–2179. https://doi.org/10.1121/1.403676.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior, 7,* 346–380. https://doi.org/10.1006/game.1994.1056.

Hofstede, G. (1980). *Culture's consequences.* Beverly Hills, CA: Sage.

Humphrey, N. K. (1976). The social function of intellect. In P. Bateson & R. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge, UK: Cambridge University Press.

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365,* 2635–2650. https://doi.org/10.1098/rstb.2010.0146.

Jensen, K., Vaish, A., & Schmidt, M. F. (2014). The emergence of human prosociality: Aligning with others through feelings, concerns, and norms. *Frontiers in Psychology, 5,* 822. https://doi.org/10.3389/fpsyg.2014.00822.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology, 80,* 557–571. https://doi.org/10.1037/0022-3514.80.4.557.

Klein, N., & Epley, N. (2015). Group discussion improves lie detection. *Proceedings of the National Academy of Sciences, 112,* 7460–7465. https://doi.org/10.1073/pnas.1504048112.

Knapp, M. L. (2008). *Lying and deception in human interaction.* Boston, MA: Pearson Education.

Knapp, M. L., & Comadena, M. E. (1979). Telling it like it isn't: A review of theory and research on deceptive communications. *Human Communication Research, 5,* 270–285. https://doi.org/10.1111/j.1468-2958.1979.tb00640.x.

Knight, C. (1991). *Blood relations: Menstruation and the origins of culture.* London, UK and New Haven, CT: Yale University Press.

Knight, C. (1998). Ritual/speech coevolution: A solution to the problem of deception. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language: Social and cognitive bases* (pp. 68–91). Cambridge, UK: Cambridge University Press.

Knight, C. (2008). Language co-evolved with the rule of law. *Mind & Society, 7,* 109–128. https://doi.org/10.1007/s11299-007-0039-1.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behaviouralecology: An evolutionary approach* (pp. 380–402). Oxford, UK: Blackwell Scientific Publications.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9,* 178–181. https://doi.org/10.1111/1467-8721.00088.

Leduc, K., Williams, S., Gomez-Garibello, C., & Talwar, V. (2017). The contributions of mental state understanding and executive functioning to preschool-aged children's lie-telling. *British Journal of Developmental Psychology, 35,* 288–302. https://doi.org/10.1111/bjdp.12163.

Lee, K. (2000). Lying as doing deceptive things with words: A speech act theoretical perspective. In J. W. Astington (Ed.), *Mind in the making* (pp. 177–196). Oxford, UK: Blackwell.

Lee, K. (2013). Little liars: Development of verbal deception in children. *Child Development Perspectives, 7,* 91–96. https://doi.org/10.1111/cdep.12023.

Lee, K., & Ross, H. J. (1997). The concept of lying in adolescents and young adults: Testing Sweetser's folkloristic model. *Merrill-Palmer Quarterly, 43*(2), 255–270.

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392. https://doi.org/10.1177/0261927X14535916.

Lyons, M., Healy, N., & Bruno, D. (2013). It takes one to know one: Relationship between lie detection and psychopathy. *Personality and Individual Differences, 55,* 676–679. https://doi.org/10.1016/j.paid.2013.05.018.

Ma, F., Evans, A. D., Liu, Y., Luo, X., & Xu, F. (2015). To lie or not to lie? The influence of parenting and theory-of-mind understanding on three-year-old children's honesty. *Journal of Moral Education, 44,* 198–212. https://doi.org/10.1080/03057240.2015.1023182.

Mascaro, O., Morin, O., & Sperber, D. (2017). Optimistic expectations about communication explain children's difficulties in hiding, lying, and mistrusting liars. *Journal of Child Language, 44,* 1041–1064. https://doi.org/10.1017/S0305000916000350.

McCornack, S. A., & Parks, M. R. (1986). Deception detection and relationship development: The other side of trust. *Annals of the International Communication Association, 9,* 377–389. https://doi.org/10.1080/23808985.1986.11678616.

McCrone, J. (1991). *The ape that spoke: Language and the evolution of the human mind*. New York, NY: William Morrow & Co.

McNally, L., & Jackson, A. L. (2013). Cooperation creates selection for tactical deception. *Proceedings of the Royal Society B, 280,* 1–7. https://doi.org/10.1098/rspb.2013.0699.

Mealy, M., Stephan, W., & Urrutia, I. C. (2007). The acceptability of lies: A comparison of Ecuadorians and Euro-Americans. *International Journal of Intercultural Relations, 31,* 689–702. https://doi.org/10.1016/j.ijintrel.2007.06.002.

Mitchell, R. W. (1986). A framework for discussing deception. In R. W. Mitchell & N. S. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 3–40). Albany: State University of New York Press.

Mitchell, R. W. (1987). A comparative-developmental approach to understanding imitation. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in ethology* (pp. 183–215). Boston, MA: Springer.

Mitchell, R. W. (1993). Animals as liars: The human face of nonhuman duplicity. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 59–89). New York, NY: Guilford Press.

Newton, P., Reddy, V., & Bull, R. (2000). Children's everyday deception and performance on false-belief tasks. *British Journal of Developmental Psychology, 18,* 297–317. https://doi.org/10.1348/026151000165706.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175–220. https://doi.org/10.1037/1089-2680.2.2.175.

Ning, S. R., & Crossman, A. M. (2007). We believe in being honest: Examining subcultural differences in the acceptability of deception. *Journal of Applied Social Psychology, 37,* 2130–2155. https://doi.org/10.1111/j.1559-1816.2007.00254.x.

Novotny, E., Carr, Z., Frank, M. G., Dietrich, S. B., Shaddock, T., Cardwell, M., & Decker, A. (2018). How people really suspect and discover lies. *Journal of Nonverbal Behavior, 42,* 41–52. https://doi.org/10.1007/s10919-017-0263-2.

O'Sullivan, M. (2003). The fundamental attribution error in detecting deceit: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin, 29,* 1316–1327. https://doi.org/10.1177/0146167203254610.

O'Sullivan, M., & Ekman, P. (2004). The wizards of deception detection. In P. A. Granhag & L. Strömwell (Eds.), *The detection of deception in forensic contexts* (pp. 269–286). Cambridge, UK: Cambridge University Press.

O'Sullivan, M., Frank, M. G., Hurley, C. M., & Tiwana, J. (2009). Police lie detection accuracy: The effect of lie scenario. *Law and Human Behavior, 33,* 530–538. https://doi.org/10.1007/s10979-008-9166-4.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5-to 10-year-old children. *Journal of Experimental Child Psychology, 39,* 437–471. https://doi.org/10.1016/0022-0965(85)90051-7.

Piaget, J. (1932/1965). *The moral judgement of the child*. London, UK: Free Press.

Premack, D. (2007). Human and animal cognition: Continuity and discontinuity. *Proceedings of the National Academy of Sciences, 104,* 13861–13867. https://doi.org/10.1073/pnas.0706147104.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1,* 515–526. https://doi.org/10.1017/S0140525X00076512.

Rauwolf, P. (2016). *Understanding the ubiquity of self-deception: The evolutionary utility of incorrect information* (Unpublished doctoral dissertation). University of Bath, UK.

Rauwolf, P., Mitchell, D., & Bryson, J. J. (2015). Value homophily benefits cooperation but motivates employing incorrect social information. *Journal of Theoretical Biology, 367,* 246–261. https://doi.org/10.1016/j.jtbi.2014.11.023.

Riggio, R. E., Tucker, J., & Throckmorton, B. (1987). Social skills and deception ability. *Personality and Social Psychology Bulletin, 13,* 568–577. https://doi.org/10.1177/0146167287134013.

Rodino-Colocino, M. (2018). Me too, #MeToo: Countering cruelty with empathy. *Communication and Critical/Cultural Studies, 15,* 96–100. https://doi.org/10.1080/14791420.2018.1435083.

Saarni, C., & Weber, H. (1999). Emotional displays and dissemblance in childhood: Implications for self presentation. In P. Philippot, R. S. Feldman, & E. Coates (Eds.), *The social context of nonverbal behavior* (pp. 71–105). Cambridge, UK: Cambridge University Press.

Santos, L. R., & Rosati, A. G. (2015). The evolutionary roots of human decision making. *Annual Review of Psychology, 66,* 321–347. https://doi.org/10.1146/annurev-psych-010814-015310.

Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology, 78,* 655–669. https://doi.org/10.1037/0022-3514.78.4.655.

Serota, K. B., & Levine, T. R. (2015). A few prolific liars: Variation in the prevalence of lying. *Journal of Language and Social Psychology, 34,* 138–157. https://doi.org/10.1177/0261927X14528804.

Shaw, J., Porter, S., & Ten Brinke, L. (2013). Catching liars: Training mental health and legal professionals to detect high-stakes lies. *The Journal of Forensic Psychiatry & Psychology, 24,* 145–159. https://doi.org/10.1080/14789949.2012.752025.

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: the scope and limits. *Trends in Cognitive Sciences, 12,* 48–53. https://doi.org/10.1016/j.tics.2007.11.008.

Sodian, B., & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *Journal of Child Psychology and Psychiatry, 33,* 591–605. https://doi.org/10.1111/j.1469-7610.1992.tb00893.x.

Sommerfeld, R. D., Krambeck, H. J., & Milinski, M. (2008). Multiple gossip statements and their effect on reputation and trustworthiness. *Proceedings of the Royal Society of London B: Biological Sciences, 275,* 2529–2536. https://doi.org/10.1098/rspb.2008.0762.

Sowden, S., Wright, G. R., Banissy, M. J., Catmur, C., & Bird, G. (2015). Transcranial current stimulation of the temporoparietal junction improves lie detection. *Current Biology, 25,* 2447–2451. https://doi.org/10.1016/j.cub.2015.08.014.

Spence, S. A., Hunter, M. D., Farrow, T. F., Green, R. D., Leung, D. H., Hughes, C. J., & Ganesan, V. (2004). A cognitive neurobiological account of deception: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society B, 359,* 1755–1762. https://doi.org/10.1098/rstb.2004.1555.

Stevens, J. R. (2008). The evolutionary biology of decision making. In C. Engel & W. Singer (Eds.), *Better than conscious? Decision making, the human mind, and implications for institutions* (pp. 285–304). Cambridge, MA: MIT Press.

Street, C. N. (2015). ALIED: Humans as adaptive lie detectors. *Journal of Applied Research in Memory and Cognition, 4,* 335–343. https://doi.org/10.1016/j.jarmac.2015.06.002.

Suddendorf, T., & Whiten, A. (2001). Mental evolution and development: Evidence for secondary representation in children, great apes, and other animals. *Psychological Bulletin, 127,* 629–650. https://doi.org/10.1037//0033-2909.127.5.629.

Sweetser, E. (1987). The definition of lie. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 43–66). Cambridge, UK: Cambridge University Press.

Sztompka, P. (1999). *Trust: A sociological theory.* Cambridge, UK: Cambridge University Press.

Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development, 26,* 436–444. https://doi.org/10.1080/01650250143000373.

Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development, 79,* 866–881. https://doi.org/10.1111/j.1467-8624.2008.01164.x.

Talwar, V., Williams, S. M., Renaud, S. J., Arruda, C., & Saykaly, C. (2016). Children's evaluations of tattles, confessions, prosocial and antisocial lies. *International Review of Pragmatics, 8,* 334–352. https://doi.org/10.1163/18773109-00802007.

Tomasello, M. (1999). Having intentions, understanding intentions, and understanding communicative intentions. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 63–75). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). In search of the uniquely human. *Behavioral and Brain Sciences, 28,* 721–727. https://doi.org/10.1017/S0140525X05540123.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation: The interdependence hypothesis. *Current Anthropology, 53,* 673–692. https://doi.org/10.1086/668207.

Towner, S. (2010). Concept of mind in non-human primates. *Bioscience Horizons, 3,* 96–104. https://doi.org/10.1093/biohorizons/hzq011.

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46,* 35–57. https://doi.org/10.1086/406755.

Trivers, R. (1985). *Social evolution.* Menlo Park, CA: Benjamin/Cummings Publishers.

Trivers, R. (1991). Deceit and self-deception: The relationship between communication and consciousness. In M. Robinson & L. Tiger (Eds.), *Man and beast revisited* (pp. 175–191). Washington, DC: Smithsonian.

Tucker, A. W. (1950). *A two-person dilemma* (Unpublished manuscript). Stanford University, Stanford, CA.

Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory, 1,* 1–30.

Von Hippel, W., & Trivers, R. (2011). Reflections on self-deception. *Behavioral and Brain Sciences, 34,* 41–56. https://doi.org/10.1017/S0140525X10003018.

Vrij, A. (2007). Deception: A social lubricant and a selfish act. In K. Fiedler (Ed.), *Frontiers of social psychology: Social communication* (pp. 309–342). New York, NY: Psychology Press.

Vrij, A., Dragt, A., & Koppelaar, L. (1992). Interviews with ethnic interviewees: Non-verbal communication errors in impression formation. *Journal of Community & Applied Social Psychology, 2,* 199–208. https://doi.org/10.1002/casp.2450020304.

Wellman, H. M. (1992). *The child's theory of mind.* Cambridge: The MIT Press.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30,* 239–277. https://doi.org/10.1016/0010-0277(88)90021-2.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75,* 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x.

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences, 11,* 233–244. https://doi.org/10.1017/S0140525X00049682.

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364,* 2417–2428. https://doi.org/10.1098/rstb.2009.0069.

Willer, R., Feinberg, M., Irwin, K., Schultz, M., & Simpson, B. (2010). The trouble with invisible men: How reputational concerns motivate generosity. In S. Hitlin & S. Vaisey (Eds.), *Handbook of the sociology of morality* (pp. 315–330). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-6896-8_17.

Williams, S., Moore, K., Crossman, A. M., & Talwar, V. (2016). The role of executive functions and theory of mind in children's prosocial lie-telling. *Journal of Experimental Child Psychology, 141,* 256–266. https://doi.org/10.1016/j.jecp.2015.08.001.

Wilson, A. E., Smith, M. D., & Ross, H. S. (2003). The nature and effects of young children's lies. *Social Development, 12,* 21–45. https://doi.org/10.1111/1467-9507.00220.

Wilson, D. S., Wilczynski, C., Wells, A., & Weiser, L. (2000). Gossip and other aspects of language as group-level adaptations. In L. Huber & C. Heyes (Eds.), *The evolution of cognition* (pp. 347–365). Cambridge: The MIT Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103–128. https://doi.org/10.1016/0010-0277(83)90004-5.

Wright, G. R., Berry, C. J., & Bird, G. (2012). "You can't kid a kidder": Association between production and detection of deception in an interactive deception task. *Frontiers in Human Neuroscience, 6,* 87. https://doi.org/10.3389/fnhum.2012.00087.

Xu, F., Bao, X., Fu, G., Talwar, V., & Lee, K. (2010). Lying and truth-telling in children: From concept to action. *Child Development, 81,* 581–596. https://doi.org/10.1111/j.1467-8624.2009.01417.x.

Xu, F., Luo, Y. C., Fu, G., & Lee, K. (2009). Children's and adults' conceptualization and evaluation of lying and truth-telling. *Infant and Child Development, 18,* 307–322. https://doi.org/10.1002/icd.631.

Culture and Deception: The Influence of Language and Societies on Lying

Darrin J. Griffin and Christian Bender

In our modern and connected society, we find that our ability to communicate and travel internationally faces little resistance across time and distance. Consequently, we are more prone to interact with others who espouse a variety of cultural values and orientations different from our own. We now find ourselves in a multicultural society. Cross-cultural interactions are inevitable, and partners in conversation will share messages that are less than accurate to what they know or how they feel for a variety of reasons. Deception is inescapable, and the literature on lying suggests that lying proliferates all societies (see Serota & Levine, 2015).

Deception is an intriguing communication construct because it influences almost every other facet of communication. We know that norms and perspectives on lying vary across cultures and that these hold implications for our cross-cultural relationships. In international business exchanges, the use of deception, even when used to avoid threatening another's face, can hold consequences for trust in future interactions. An example provided by Kam, Kim, and Sharkey (2008) unravels this cross-cultural dynamic—if a Japanese businessperson provides information to an American counterpart

D. J. Griffin (✉)
Department of Communication Studies, The University of Alabama,
Tuscaloosa, AL, USA
e-mail: djgriffin1@ua.edu

C. Bender
Department of Modern Languages and Classics, The University of Alabama,
Tuscaloosa, AL, USA
e-mail: cbender@crimson.ua.edu

that is face-saving, but is perceived as less than honest by the recipient, then this might be perceived as a negative counter by the American and create a loss of trust for future interactions. Kam et al. alternatively illustrate that the American's blunt and direct expectations for feedback might be received as inappropriate by the Japanese counterpart, who may in turn expect that messages are provided in less face-threatening ways to maintain harmony across relationships. Dialectics in the way *truth* is used, as illustrated in the example above, often create tensions during intercultural communication.

Communicating truthfully across a variety of situations can challenge even the most skilled speakers. Communicators in virtually every interaction must decide both *what to say* and *how to say it* (Knapp, McGlone, Griffin, & Earnest, 2016). Expressing our thoughts, perspectives, and views on reality, all while maintaining our apparently honest nature requires social skills that we learn throughout our lives. The influence of culture is ever present as we hone our communication skills. As the world becomes increasingly more connected, the influence of cross-cultural communication, the growing need for intercultural competence (Holliday, 2016), and a better understanding of the influence that culture plays in perspectives on deception are vital.

While the majority of deception research has taken place in the US and Western Europe (see Levine, Ali, Dean, Abdulla, & Garcia-Ruano, 2016), this chapter will examine studies that explore a variety of cultures and how they interact with and influence deception. This chapter provides readers with an overview of what we know about cultural perspectives on lying and deception and concludes with suggestions for topics of inquiry that readers might consider using in their own research. We start by extracting a brief overview into the meaning of culture, provide an overview of the dimensions of cultures (e.g., collectivism and individualism), and then describe why deception pertains to various cultures guided by principles and practices that tend to differ from one another. Next, we explore how children are taught about truth-telling and lying across different cultures. We will visit the research that studies lie detection across cultures and conclude with recommendations for future research in this timely area of deceptive communication.

DEFINING CULTURE

Though culture is often taken for granted, it deserves to be (and certainly has been by scholars) more critically examined. Whereas every person has an identity, culture is akin to the identity of a society, its lifeblood. It consists of the customs, beliefs, art, language, behaviors, food forms, and communication habits found among a group of people. Fortman and Giles (2006) remark, "Culture is ubiquitous, multidimensional, and complex" (p. 91). Defining culture is fraught with difficulty because it is a term expressed with great variability, and perhaps disagreement, among academics and practitioners. Baldwin, Faulkner, and Hecht (2006) provide an extensive overview

of the historical transformations on definitions of culture. In over 85 pages of their volume, they provide a list of hundreds of definitions for culture employed by scholars of varying disciplines. Scanning their exhaustive list to choose one single definition, we relied on two criteria: (1) parsimony and (2) recency. Relying on these two criteria (note the irony that the two heuristics are those enculturated by our empirical training), of the dozen definitions that use approximately 15–20 words to define culture, the most recently published from this subsample explains that "culture is defined…as a historically shared system of symbolic resources through which we make our world meaningful" (Hall, Covarrubias, & Kirschbaum, 2002, p. 4). We have found that the majority of definitions of culture can be synthesized or expanded to include at least some notion of Hall et al.'s (2002) definition, which includes something people do, share, and the inseparability of communication in culture (see Cargile, Giles, & Clément, 1995). Triandis (2011) provides an excellent overview of culture (and a definition) and an illustration in the way of metaphor, whereas culture is akin to a society's memory.

The purpose and functions of culture is a complicated matter; although there are many authors who provide various predicted purposes for culture, we believe that the function of culture is essentially twofold. First, it provides a bond among the people who share it and an intrinsic motivation for taking care of one another and their tangible and valuable resources. Second, cultural indicators allow humans to brand themselves to others outside of their sphere. Thus, culture exists only with borders. That is, culture only comes into play when there is another entity, similar in design and function but starkly different in execution. To understand a culture, there must be a contrast by which to compare the differences among them (Fortman & Giles, 2006). In other words, understanding a particular culture relies on at least one other culture for a point of comparison. To understand culture, we must look both within a culture and also across cultures. In viewing culture in these ways, we can then explore how cultures can impact deception.

A culture itself is a collective identity in that it is less associated with one's personal identity than with a group or societal identity (Brewer & Gardner, 1996). Culture is a by-product of the necessary human condition of living in groups (Graham, Meindl, Beall, Johnson, & Zang, 2016). Humans band together in groups for safety, and these groups must have shared norms to survive and flourish. Though these norms and customs comprise a group's culture, most cultures are not simple homogenous happenings. For example, in the US, different states follow and enforce laws with a great deal of variance (Graham et al., 2016). Another example lies within the Deaf cultural community. In it, there are intersections across various other identities such as Black Deaf, who have their own dialects of sign language (known as Black ASL; McCaskill, Lucas, Bayley, & Hill, 2011) and Deaf LGBTQA+ groups who form their own social and advocacy groups (see Blau, 2017). The intersectionality of identities within cultural groups makes them diverse and complicated happenings.

As noted, culture does not necessarily shift at the ethnic or geopolitical boundaries of nation-states, nor does culture begin or end with the social constructs of language or political parties. However, while culture is often assumed to exist at the large macro-level classifications made up of countries and national boundaries, it also exists at the micro-level. Two such examples of micro-level cultures are corporations and family units. Corporations, though smaller than nations, still create and enforce unique rules and practices that add to the commonly held idea of *corporate culture*. Family units may craft their own unique culture in fashioning rituals, customs, or superstitions. The expression that *it has always been done that way* may explain the rules that create these cultures. First-year college students take on the long-held traditions and lore of their new campus as if these mechanisms have always dictated their own ideologies. Singing the "Rammer Jammer Cheer" after a football game win is but one of the many traditions carried on by members of The University of Alabama community. The comedian and actor Jim Carrey recently captured the essence of how our cultural identities are but "abstract structures that hold us together" (Fitzgerald, Gabai, Jonze, & Smith, 2017).

Corporate trainers teach the notion of culture by illustrating an anecdote involving chimpanzees and the concept of learned helplessness. As the narrative is sometimes told, the primates climb a ladder in their cage that leads to food, but are sprayed with a water hose in doing so. The chimps learn to avoid the ladder for fear of being sprayed. Any new chimps introduced to the cage that attempt to climb the ladder are swiftly attacked by the other, more learned members of their community, and thus become more entrenched in the troop's culture. As the story goes, this social learning is carried across a few generations of chimps, and their avoidant behavior of the ladder is carried on long after the punishment is executed. In time, all of the chimps that were punished no longer exist, and the cruel keeper who enacted the punishment is also no longer present—nonetheless, the ladder is avoided because of learned cultural norms. This narrative aims to parallel corporate culture in how learned helplessness can be passed on in irrational ways because *things are just done this way*. Interestingly, this story seems to be based on the research of Stephenson (1967) who conducted an experiment on culture and learning among Rhesus monkeys in this very fashion, but used airbursts as the negative stimulus.

Almost all cultures consider deceptive acts with harmful consequences to be a violation of expectations (Levine et al., 2000). However, when and how to use deception as a communication tactic is a matter that differs across cultures (Lapinski & Levine, 2000). One example of this differentiation occurs when Russians and Americans rely on deception to support an in-group member who is failing when pitted against an outsider in competition. When faced with a hypothetical scenario that puts an underperforming in-group member against a successful outside member, Russians tended to respond

with deceptive excuses while Americans relied more so on honest communication (Bessarabova, 2014).

Norms, practices, and customs of cultures are shared with members via communication. While communication is one mechanism used to express culture, it is also something that is heavily influenced by the expectations (e.g., norms) of different cultures (see Lapinski & Levine, 2000). All communication entails decisions, and often these decisions encompass the degree to which one is forthcoming with information and may lead to choices involving truth-telling and deception. However, cultural orientations may play a serious role in these decisions in that the orientation of a culture in which one finds themselves drives the methods, motives, and views of deceptive acts.

Cultural Orientations/Dimensions

As mentioned, culture is not a simple construct. The many dimensions and orientations of culture separate social networks from one another based on the dynamics that members enact when engaging with one another. Hofstede (1984) coined what is known as the dimension of collectivism and individualism, and Triandis (2004) added two more dimensions, complexity, and tightness, all of which have been used to examine the nuances of culture as it moderates other variables across disciplines for several decades. Below, we offer a summary of each of these dimensions and review relevant cross-cultural studies of deception.

Cultures have complexity (Triandis & Suh, 2002); they range from least complex, such as hunter-gatherers, to very complex, like information societies (Triandis, 1989). Complexity equates to the number of cogs turning within the engine that is culture. A culture that has many factors complicating its social construction, such as numerous cities, modes of transportation, taxes, and religious belief systems, are considered more complex than cultures operating with fewer of these social constructions. For comparison, consider the roaming bands that existed approximately 15,000 years ago and consisted of roughly 30 members, to the millions residing in the modern metropolises of today. Of course, the former is far less complex than the latter (Triandis, 1995). Further, the complexities belonging to a particular culture may denote what specific things the culture values. For example, there are over 600 terms related to communication about *camels* in Arabic and dozens to relate to the concept of *car* in English (see Triandis, 1995), illustrating the focus of these cultures on particular forms of transportation, but this is only one mechanism of their cultural complexity.

A second cultural orientation is tightness, which refers to how rigidly a culture practices its norms. Triandis (2004) created a cultural tightness scale to measure the variability in which tighter cultures staunchly observe their rules and customs, and looser cultures more often deviate from such customs with little reproach. Gelfand, Nishii, and Raver (2006) theorize that a culture's

level of tightness may depend on how the culture is developed. Through comparing hunter-gatherer cultures to those with more agricultural values, they illustrate how the concept of tightness might be influenced by the social structures of societies. The former is designed such that individuals must be creative in how they obtain and hunt for food, while the latter requires a strict system of rules to successfully grow crops. Tightness can be observed within modern cultural business practices. For example, a legal firm conducts business in a case-by-case basis, whereas a telecommunications company may develop and implement their business practices through rigid rules that are enforced across all scenarios involving customers (e.g., billing structures for mobile phone plans). The tightness of culture influences the rigidity of rules that govern the behavior of its members.

The last cultural orientation explored in this chapter is collectivism-individualism. Cultures are identified on this dimension in how the individuals within these groups pursue their goals. In individualist cultures, individuals place importance on their personal goals over the objectives of their group. Conversely, individuals within collectivist cultures either do not discern between personal and group goals, or raise their group's intentions over their own (Triandis, 1995). Group interdependence is another facet that dictates whether a culture is collectivist or individualist. Individualist cultures display far less group cohesion than collective cultures—persons hailing from collectivist cultures are generally more entrenched in the lives of other members from their groups and more tightly follow their group's norms and customs.

While the theory of dimensions/orientations dominates communication research and studies exploring cultural factors that influence human behavior, recent studies are considering other cross-cultural factors that mediate relationships between communication and culture. For example, a recent study evaluates culture from the standpoint of religion, monogamous marriage, and kinship systems (Henrich, 2015). But, the vast amount of work conducted on culture does include Hofstede's cultural dimensions, and readers must not ignore the contribution to the literature these works have made.

Cultural Orientations and Deception Research. Researchers have examined how cultural orientations influence duplicity across cultures (e.g., Seiter, Bruschke, & Bai, 2002). Collectivist cultures promote interconnected communities whose members reverentially follow agreed customs and therefore value harmony as a group. Some might expect these members to employ deception of lower stakes (i.e., white lies) to promote social cohesion/ harmony. These cultures tend to promote group well-being, and so individuals representing these cultures would also be more prone to lie for the sake of the group (i.e., blue lies; discussed below when examining stakes). Alternatively, those from individualistic cultures tend to yield highly autonomous goals, and, therefore, should be expected to lie in order to benefit themselves more often than others. A culture's dimensions may dictate what

a lie is depending on who benefits from the potentially deceitful situation. As Bond and Atoum (2000) have discovered, a tighter culture equates to its members more staunchly obeying its rules, and thus the acceptability of a lie is generally received more harshly by others than in looser cultures. There are few studies solely evaluating cultural deceit due to cultural complexity and tightness/looseness. However, Triandis (2004) cites specific examples regarding these two dimensions. First, he describes the Taliban in Afghanistan, who execute people for relatively simple offenses (e.g., music choice) as an extremely tight culture. Second, he writes that Thailand is a loose culture, in which there are seldom any rules binding the majority of the people. From these examples, it is evident that a culture's tightness will certainly relate to their views, behaviors, and punishments of deception.

In the last decade, an upsurge of research in deception has examined the influence that cultural dimensions may play in lying. In one study, individualist (i.e., American) and collectivist (i.e., Korean) participants judged an ambiguous statement on its deceptiveness. Seventy percent of the American participants judged a particular message to be deceptive, compared to the 35% of Korean participants who perceived the message in this way (Park & Ahn, 2007). More recently, studies focus on intercultural deceptive acts themselves, and the motives behind them, rather than on the intercultural judgment of lying. For example, researchers posed a hypothetical scenario displaying an underperforming in-group participant against a successful outside participant. The in-group could either lie and cover up their member's failure or disclose the truth. In this study, Russian participants (collectivists) more often responded with deception while US participants (individualists) responded more often with the truth (Bessarabova, 2014). In a separate but related study, Korean participants (complex collectivists) and American participants (complex individualists) considered the act of lying for the benefit of a friend (Choi, Park, & Oh, 2011). Findings revealed that the Koreans viewed lying for a friend more positively, while the Americans perceived the use of truth in a more positive manner. Based on the findings, the researchers posited that collectivists consider maintaining group harmony (e.g., lying for a friend) more important than maintaining one's integrity (e.g., telling the truth) and that individualists consider the opposite to be true.

Though cultural orientations may help researchers study cross-cultural deception, some argue that the orientations need to be further refined; that is, cultural dimensions are measured differently among researchers and the concepts may be too broad to precisely categorize a culture. Specifically, Uleman, Rhee, Bardoliwalla, Semin, and Toyama (2000) argue that the orientations used by many scholars fail to distinguish dimensions that exist *within* a culture. Levine et al. (2000) explain that cultural orientations are unable to detect cross-cultural deception altogether. They argue that, though some research may detect deception when individuals deviate from their culture's prescribed dimension, these deviations do not automatically equate to

an act of deception. However, when individuals deviate from normative cultural behaviors, this does tend to influence deception detection judgments. Those who stray from their culture's norms were judged to be less honest than those who acted within those norms (Levine et al., 2000). Despite recent critiques, research continues to employ the dimensions as a way to make comparisons related to deception across cultures.

Cultural orientations affect a culture's morality (Graham, Meindl, Beall, Johnson, & Zhang, 2016). Simple, loose, and collectivist cultures tend to promote a morality that dictates the duties one owes to one's community. Not surprisingly, members of an organization who come from differing cultures must obtain cultural competence in order for the organization to thrive (Triandis, 2002). In terms of future studies, Choi et al. (2011) suggest that researchers focus on how deceivers explain their motives for lying, and how a particular audience affects the deceiver's delivery of such motives. Perhaps most importantly, as suggested by Taylor, Larner, Conchie, and Menacere (2017), researchers might consider focusing on the unique differences of the individual, the group, and the culture as a whole, as they explore individually construed deceit.

Face, Self-construal, Directness

The work of Goffman (1967) introduces the notion that one has a *face* or participates in *face-work* in their everyday social interactions with others. Face, according to Goffman (1955) "is defined, as the positive social value a person effectively claims for [them]self by the line others assume [they] have taken during a particular contact" (p. 213). A line is a communicative act whereby one affirms some interpretation or view of a context or situation. In other words, face is merely a way of seeing oneself in a positive light among the many social relationships and dynamics in which they find themselves existing. One may wonder (as the first author does) whether they are a fair and loving husband or simultaneously wonder whether their hand-crafted table will turn out sturdy, and consider whether others like the rustic design chosen which shows the patina of the century-old wood used to surface the table. These are but simple illustrations of the infinite other situations where one may communicate verbally and nonverbally to negotiate their value with others. These examples illustrate personal narratives in which the ego factors as a part of their identity, but many other cultural dynamics direct how interactional partners might engage in face-work with one another in real time or across their relationships or group memberships.

Conflict is a context rife with opportunities for infringing on the identities and face concerns of others (e.g., Oetzel & Ting-Toomey, 2003). As with most conflict, there is often a perceived notion that one's goals must be sacrificed for the well-being of others. Conflict often entails that the participants or combatants must be less than honest in their response, due to the power dynamics, cultural norms, or relational history of those engaged in these

matters. Borrowing from one of our earlier examples, a Japanese businessperson may be hesitant to provide information (or feedback) that will be received as an insult or which might infringe on the other's view of themselves as a successful partner in the exchange—in this case, business relations. Whereas, the businessperson from the US considers it their duty to be honest and direct with others and give them (their own perceived) truthful appraisals of the situation, regardless of how damning or embarrassing the information might be to the recipient. However, a person from a culture that values honesty over face concerns may be more likely to judge a less than forthcoming counterpart as deceptive when the goal of social harmony is chosen over directness (see Lapinski & Levine, 2000).

Ting-Toomey, Oetzel, and Yee-Jung (2001) found a person's self-construal (see Gudykunst et al., 1996), or the way in which one identifies with the self versus their larger culture group, to be effective predictors of conflict styles. Accordingly, the work of Lapinksi and Levine (2000) reveals how independent and interdependent self-construals manifest differences in the perception and judgment of the veracity of messages. The cultural landscape, or the intercultural situation, where individuals communicate certainly impacts the content and development of their messages and how those messages are judged. Two factors may affect how one navigates and negotiates their face: One, the norms included in a cultural landscape which dictate whether messages ought to be direct or indirect, and two, how closely embedded within a cultural landscape someone is, or thinks they are. While cultural dynamics, orientations, and self-construal play a role in the nuanced ways people lie and how deceptions are judged, general underlying motivations for falsifying the beliefs held by others are vast.

Motivations for Lying: Why People Lie

Generally, people lie to avoid punishment or to obtain a reward which they cannot obtain using the truth. However, truth-telling is a heuristic used to achieve communicative goals (exceptions are found in pathological liars and among those with personality disorders). There are many reasons people justify their use of deception. Motivations for lying range from benefiting financially or socially, or to increase one's own perception of themselves or another person. Not all lies are told for the self; some can be told for altruistic reasons as well. When lies are told, though, there is often a benefactor. People do not lie as a mean in and of itself. They lie to achieve a motive or goal. Those who benefit from the consequences of a lie may be the equivocator themselves, another person, both parties in a relationship, or those belonging to a larger social entity.

When telling the truth, our default communication strategy (see Levine, 2014), might present a problem for goal attainment then deception reveals itself as an attractive alternative strategy. The work of Ekman (1997) reveals the following motives for deception: (a) avoiding punishment, (b) obtaining

rewards, (c) protection of others, (d) self-preservation, (e) admiration/pride, (f) avoiding undesirable social interactions (e.g., awkwardness/embarrassment), and to maintain (g) privacy and (h) power. Other motivations have been further categorized and organized based on factors such as the beneficiary of the lie (Hample, 1980; Metts, 1989) or rewards based on different social situations (Camden, Motley, & Wilson, 1984). Overall, the research on motivations for deception from an intercultural communication perspective is sparse, but has been explored via multicultural motivations for lies and views of acceptability across the various relationships, benefactors, and motives of cross-cultural deception (Seiter, Bruschke, & Bai, 2002).

Most recently, Levine, Ali, Dean, Abdulla, and Garcia-Ruano (2016) explored the motives for deception across four non-Western cultures and a US sample. Relying on truth-default theory as a lens, and the proposal that "the motives that prompt deception are pan-cultural, but the situations in which those motives become salient and obstructed by the truth are culturally variable" (p. 4), the majority of lies reported by Egyptians, Guatemalans, Pakistanis, and Saudis were motivated by self-serving purposes such as economic gain; lies to avoid another person also emerged as a dominant theme. Overall, the predictions of truth-default theory hold that liars rely on falsity when the truth presents a problem in achieving their goals (see McCornack, Morrison, Paik, Wisner, & Zhu, 2014).

A Pan-Cultural Lie Motive Typology

Relying on the tenets of truth-default theory (TDT; Levine, 2014), that is, that those who judge others tend to perceive them as honest the majority of the time (independent of actual veracity), Levine et al. (2016) recruited participants from Egypt, Guatemala, Pakistan, Saudi Arabia, and the US to recall an instance of deception that occurred to them (directed to them by another person) in the recent past. Using open-ended responses, the authors generated a list of motives for deception that occurred across these countries. This procedure was used to develop the following list of ten pan-cultural deception motives:

1. personal transgression; lie intended to conceal a wrong
2. economic advantage; lie intended to accrue wealth
3. non-monetary personal advantage; lie intended to benefit the liar (non-monetary)
4. social-politeness; lie intended to uphold a societal standard or to avoid being rude
5. altruistic lying; statement intended to protect or benefit someone other than the liar
6. self-impression management; lie intended to bolster one's appearance
7. maliciousness; lie intended to harm another

8. humor; harmless deceit intended to joke, play, entertain, or induce laughter
9. pathological lying; delusions, disregards for norms, compulsive false statements
10. avoidance; acts intended to avert another from the truth.

Overall, these findings affirm the notion supported by TDT that liars enact deceptive communication when telling the truth presents a problem for obtaining goals. Additionally, it supports the assumption that there may be universal motives for deception across cultures. This attempt to situate lie motives cross-culturally is a novel and recent endeavor, but promises future inquiries into this type of research. Exploring lie motives across cultures will assist us in understanding the reasons people decide to tell lies. It is clear that people tend to tell different types of lies based on the various outcomes of situations.

Stakes

White lies are those deceptions which are relatively inconsequential and possess little threat to either the liar or their target of their statement if discovered. These low-stakes deceptions occur frequently in social conversation (Knapp & Comadena, 1979) and often revolve around one's feelings about something mundane, such as the enjoyment of a social gathering (e.g., "this was so much fun, let's do it again soon"), or the preference for the flavor of someone's sweet tea. Many of these interpersonal lies are told to interact effectively with a more powerful person (such as a boss or a parent), save face, avoid hurt feelings, or dodge friendly requests (Lippard, 1988). Because norms and expectations for such interactions vary across cultures, research has indicated that views of acceptability of these types of social lies differ across cultural groups (Seiter, Brushke, & Bai, 2002).

In popular culture, the word *lie* is often used to illustrate a more serious communicative behavior—one that has major penalties such as loss of money, breach of trust, or damage to property or life. These consequential high-stakes lies, or what have been termed *black lies* (see Camden, Motley, & Wilson, 1984; note however, that this terminology has not been adopted by others to describe serious lies), are statements or actions which mislead others in ways that may cause physical or psychological damage or alter one's sense of reality in an extreme way (e.g., gaslighting; Abramson, 2014). These transgressions often involve gains or losses that go beyond those of white lies, that commonly serve as a social lubricant. While they can be enacted during interpersonal relations, they are also told frequently by organizations, institutions, political parties, and governments. These lies may exude the highest levels of stakes because they have been associated with very large monetary gains and extreme loss of life (e.g., war lies told by government officials). While political

lies may be told to protect the accused, or to prevent a group of people from a previous mistake, they may also be wielded as propaganda intended to benefit the masses from an outside threat such as an enemy country or a political opponent.

Lies told to protect an in-group have been referred to as *blue lies* (Klockars, 1984). The term originally described the ways in which police and law enforcements must at times rely on deception to uphold the law, to protect their fellow officers, or to obtain information from criminals. Barnes (1994) extended the term's definition in his research on the motivations for lying in cultural groups, stating that blue lies protect other cultural insiders from consequences which may otherwise be enforced by telling of the truth (see Lee, 2013).

Researchers have not yet conceptualized whether blue lies extend to both low- and high-stakes deceptions, but instead describe them as deceptions which are told in a particular context (i.e., in- vs. out-group situations). For example, blue lies would promote group well-being and aim to decrease the advantage of a target who is not a cultural member. More recently, the concept of blue lies has been applied to the political behaviors of politicians and the perception of their misrepresentations based on recipients' partisan affiliations (Flynn, Nyhan, & Reifler, 2017). Certainly, we could argue that there is a lively political culture in the US, in that political partisanship or international political affiliations have some bearing on the views and behaviors of people. Consequently, as we illustrate throughout this chapter, this *culture* would influence people's uses and views of deception. While this research is in its infancy, we are likely to see an upsurge in research examining how our post-truth climate will provide a great deal of fodder to both deception researchers and those interested in cross-cultural lying.

FREQUENCY OF LYING

Everybody lies. Though this is something that is hardly shocking, the degree to which people lie is a question that recent research has explored with more focus than in the past. An original inquiry into the frequency of lie-telling in everyday life by DePaulo, Kashy, Kirkendol, Wyer, and Epstein (1996) found that on average people report telling lies once to twice a day. More recent research, however, has shown that these reported averages on daily lie-telling are skewed and misrepresented, and that lie behavior is not evenly distributed across populations (Serota & Levine, 2015; Serota, Levine, & Boster, 2010). The few prolific liars who tell several lies per day carry the majority share of digressions on speaking the truth and inflate the reported averages on use of deception. Recent research (Serota & Levine, 2015) continues to fine-tune what we know about the prevalence of lying, and there is continued international support that most people do not lie very often (excluding white lies). There are a few prolific liars who also cheat or exhibit psychopathic tendencies frequently (Halevy, Shalvi, & Verschuere, 2014).

Learning to Tell Lies, Children Studies

Studies involving childhood deception suggest that even those as young as three years of age have learned that it is appropriate to tell white lies to avoid embarrassing another (e.g., based on others' appearances). Young children can also communicate these types of falsehoods in such a way that adult judges are not very accurate at discerning their messages from those of truthful children (Talwar & Lee, 2002). Theory of mind, or the ability developed by the age of three to understand that others do not share one's own vantage point, enables children to lie (Ding, Wellman, Wang, Fu, & Lee 2015). However, by the age of four to five, children can communicate intentionally misleading information to establish a false belief in others. By the age of seven, most children show a fully developed theory of mind and advanced lying skills (see Talwar, Crossman, & Wyman 2017). There is recent evidence that culture may influence children's theory of mind and false-belief recognition (Duh et al., 2016).

Lee, Cameron, Xu, and Board (1997) found differences between Chinese and Canadian children in the age groups of seven, nine, and eleven years of age in their evaluation of lies told in pro-social (politeness) versus antisocial (misdeed) situations. Overall, their findings revealed that Chinese cultural norms for modesty and self-effacement were viewed more for pro-social reasons than for those of the Canadian children. Again, using a very similar format for examining the views and categorizations of lies and truths among the same two cultural groups of children, Fu, Xu, Cameron, Heyman, and Lee (2007) found Chinese children more likely than Canadians to endorse lies that benefit the group (e.g., blue lies; Lee, 2013), and rate individualistic lies more negatively. Additionally, unlike American children, Korean children tend to regard lying for a friend less negatively (Choi, Park, & Oh, 2011). Cultures may also socialize their members and promote deception to protect in-group vs. out-group members. As mentioned previously, Russians, as a cultural group, are more prone to deceive than Americans in performance tasks when pitted against insiders versus cultural outsiders (Bessarabova, 2014). This finding shows that cultures can both condemn lying to one's own group and simultaneously praise lying to opponents (see Dunbar, Jensen, Harvell-Bowman, Kelley, & Burgoon, 2017).

Overall, what the aforementioned studies indicate is that the cultural socialization of children seems to dictate how children (and adults) learn about morals, ethics, judgments, and the degree to which certain types of behaviors should be thought of in terms of lying and truth-telling. Decisions related to the use of truth or deceit are driven by the target's identity and classification as an in- or out-group member. The influence of culture on deception in child development is an area that has been well examined, but future research is certainly warranted as there are many mechanisms responsible for the influence of deceit.

Media Choice for Deception

An early cross-cultural study examining Korean and American participants' tendencies to rely on face-to-face communication versus those of mediated channels of lying found preferences across both cultures for more synchronous and less recordable media for deception (Lewis & George, 2008). These findings support the earlier findings of Hancock, Thom-Santelli, and Ritchie (2004) that showed that communicators tend to lie via rich channels and those less prone to being captured for later analysis (e.g., an email can be shared or reexamined while a conversation cannot). However, it is important to note that people make media choices for many reasons, including factors indirectly related to deception (e.g., avoiding punishment). In the case of pro-social lies (i.e., white lies), saving face is but one culturally driven factor that may influence media choice when employing deception (Furner & George, 2014). Recent research which examines cross-cultural tendencies for online deception has found, unsurprisingly, that with the rise of the internet, people across cultures report telling more lies via the internet than in face-to-face interaction (Marett et al., 2017). However, as with other studies of intercultural deception, they also found that preferences and situations for deception—even online—vary based on espoused cultural values.

LIE DETECTION

As this chapter focuses on deception and culture, we ought to explore the research on lie detection across cultural groups. Overall, the majority of cross-cultural deception research focuses on the detection skills of different cultural groups and extends theoretical tests administered to cross-cultural populations designed to support previous findings in the literature (see Park & Levine, 2017).

Bond, Omar, Mahmound, and Bonser (1990) tout that they were the first to conduct a cross-cultural study relying on nonverbal behavior as a cue to deception. While they may have been the first to examine international cultures in nonverbal deception detection, Berger (1977) had previously relied on a sample of Deaf student participants attending Gallaudet College (now Gallaudet University) to compare hearing participants in a decoding task of liars recorded on film. While he did not frame his group comparison as one between cross-cultural groups, but rather justified the design of his research based on the visual nature of nonverbal lie detection (that is, using a visual detection group, the Deaf cultural community), there is ample support that there is a Deaf cultural/linguistic community in the US (Nomeland & Nomeland, 2016). In this way, some of our own research has explored how the beliefs and communicative processes of the Deaf cultural community may influence their judgments of deception (Griffin & Frank, 2018).

Intergroup bias (Hewstone, Rubin, & Willis, 2002) dictates that people view those perceived as part of their in-group as more favorable than those

outside of their circles (albeit these groups can be large, e.g., national affiliations). Likewise, people perceived as being more similar in a characteristic (i.e., in-group members) are generally seen as more trustworthy, or are given a benefit of doubt when lying might otherwise be suspected of them (Castillo, 2015). Truth-default theory (Levine, 2014) predicts that most people view others as honest despite the actual veracity of their messages. Thus, there is a *truth bias* at play that allows us to cooperate with one another.

Cultural differences in communication styles, norms, or expectations (e.g., looking/not looking someone in the eye when speaking to them) can lead to differences in beliefs related to honesty (Dunbar et al., 2016). These may also include verbal indicators such as slang and accents, but these differences also exist in many nonverbal differences ranging from the use of facial expressions (or suppression of them, e.g., Japanese norms of expression in public) or clothing and garb used by some groups (e.g., concealing the face or head with a niqab, hijab, burka, or other head garments worn by some Muslim women across the globe). Generally, a lack of familiarity with the customs, language, and communicative acts of those belonging to another cultural group may raise interest and suspicion, and beliefs in skepticism may be increased accordingly.

Bond and Atoum (2000) examined the lie detection abilities of Americans, Jordanians, and Indian nationals and did not find that viewing the messages of a speaker from another culture raised suspicion or lie rates. However, the authors did find that there were some lingo-centrism effects whereby members of other linguistic groups from the same geographic cultural region were judged as more deceptive. Accordingly, there is research that supports the phenomenon that speakers lying in their second language may be prone to being detected (Cheng & Broadhurst, 2005), but these second-language speakers may also be judged erroneously as liars when they are telling the truth. Research does not support that certain cultural groups maintain better detection skills than their counterparts. However, recent research by Park and Levine (2017) has indicated a robust finding associated with the truth-default theory that proportions of truth/lie statements will correlate with detection rates—even in a cross-cultural sample.

Beliefs in Cues to Deception

The Global Deception Research Team (2006) has conducted the largest cross-cultural deception study to date, in which they examined believed cues to deception across 75 different countries to illustrate the dominant pan-cultural stereotype that liars avert their gaze. The majority of deception researchers agree that there is no substantial evidence that this cue will betray a liar, but rather that social and pop-cultural influences (Hurley, Griffin, & Stefanone, 2014) prolong the belief that this (inaccurate) cue (and others) correlates with actual lying behavior. Actual versus believed cues to deception

differ drastically (Feeley & Young, 2000), and the last decade of deception research has shifted the paradigm from nonverbal cues to those that are captured through the content of verbal messages and compared against the context of what is known or happening when a lie is told (Blair, Levine, & Shaw, 2010).

In an experimental design of a lie detection activity (e.g., bluffing game), Holm and Kawagoe (2010) compare Swedish and Japanese participants in their self-reported lie detection beliefs because "many subjects both in Sweden and Japan express moderate to strong beliefs in their lie detection ability" (p. 311). Relying on the notion that cultures differ in their risk-taking behaviors, the authors examined differences between these groups. Despite finding no significant differences in lie detection ability, there was a significant drop in confidence across both cultural groups in lie detection skills when monetary loss was at stake in an activity. In other words, people are more likely to believe that they excel at lie detection until stakes are involved. The majority of lie detection studies examine individuals' abilities in detection and then, through cross-cultural comparisons, discern where they identify or belong. Perhaps, relying on the recommendations of Frank, Feeley, Paolantonio, and Servoss (2004) is warranted, whereas small groups may show increases in lie detection abilities because they possess a collective knowledge of communicative norms—even those about truth-telling—in a more nuanced way than individuals.

Overall, there is insufficient evidence that any cultural or linguistic group is superior at lie detection as compared to any other. On the contrary, research on the acceptability of lies provides ample support that cultural orientations or views influence how lies are judged, not against truths, but against moral contexts. While deception research has explored differences in lie detection ability and lie acceptability beliefs dawning with the work of Bond et al. (1990), or perhaps as far back as Berger's (1977) analysis of hearing and Deaf cultural participants, we are left examining current trends in deception based on the research paradigms of scholars who are currently active in the field.

Recommendations for Future Research

Cross-cultural research continually discovers new phenomena and extends our current theories and paradigms related to deception. There are several areas of active deception research ripe for examining both cultural and linguistic differences among groups. For example, online environments which share web-based platforms such as discussion forums (e.g., Reddit), streaming video gaming, virtual monetary exchanges (e.g., cryptocurrency), and online marketplaces such as Amazon, Craigslist, or Etsy are but a few of the contexts where mediated communication and international cultures collide. Extending the work of researchers who examine text-based language (e.g., Markowitz & Hancock, 2016) to cultural contexts will explore if language betrays deceptive

intentions across cultures and languages. As shown from the citations in this current chapter, the ongoing research of Levine (2017a) that addresses method execution, theory production (Levine, Blair, & Carpenter, 2017), and applied uses of deception detection (see Levine, 2017b) has already been applied to cross-cultural contexts. However, scholars interested in conducting deception research in this domain should familiarize themselves with both his work on deceptive language and the theoretical work of his colleagues (e.g., Clementson, 2017; McCornack, 2015).

The recent research of Dunbar et al. (2017) and ten Brinke et al. (2017) reveals that indirect measures of verbal and nonverbal behavior may increase detection rates of liars. These applications aid in analyzing cross-cultural groups in that as communication occurs internationally, there are many times when communicators speaking different languages are left with only a portion of others' intentions and future behaviors. Additionally, and as Dunbar (2018) has recommended, future researchers should study the moral perspectives of cultures and how this might influence telling lies to specific individuals in certain contexts, and whether the detection of deception may also be viewed as appropriate—depending on the cultural orientation or region with which the deceiver is associated.

A current paradigm for deception detection relies on active questioning (Blair et al., 2010), that is, examining what is said as it compares to what is happening around statements. This technique for detection is of grave importance to the critical analysis of cross-cultural deception. The biases that exist among members of one culture may cause them to be critiqued by those of another culture during the telling and detection of the lie. How North and South Koreans feel about one another, or how those from European countries under the threat of the Kremlin might feel toward Russians, is of major importance as we aim to understand the many cultural variables associated with deception. Also, a lack of cultural competency in a dominant group as they aim to understand (or detect deception in) another group is something that would certainly influence the outcome of effective communication attempts, including lie catching.

While deceptive communication researchers aim far and wide to examine (in)consistencies in the theories and practice of the sub-field, they ought not to forget the many intersectionalities and identities that exist within mainstream US populations (among the participants recruited for domestic studies). Sub-cultural groups are prolific and range from small language minorities (e.g., the Deaf community who use American Sign Language), to those who identify based on their gender or sexual preferences (i.e., LGBTQA+ communities), and small cliques of diverse people who organize themselves around an activity (e.g., gangs or bikers—or biker gangs).

There is much to gain by realizing that culture is a construct that, while useful for macro applications, can also be wielded to understand differences in communication of a variety of groups. These groups, ranging from smaller

cliques, such as local and widespread police culture, college culture, and perhaps even the culture of academicians, to large and dense nation-states, contain cultural paradigms. These paradigms, in turn, sometimes circulate within their own cultural networks and often fail to be shared by others beyond their subfields.

References

Abramson, K. (2014). Turning up the lights on gaslighting. *Philosophical Perspectives, 28,* 1–30. https://doi.org/10.1111/phpe.12046.

Baldwin, J. R., Faulkner, S. L., Hecht, M. L., & Lindsley, S. L. (Eds.). (2006). *Redefining culture: Perspectives across the disciplines.* Mahwah, NJ: Routledge.

Barnes, J. (1994). *A pack of lies: Towards a sociology of lying.* Cambridge: Cambridge University Press.

Berger, R. E. (1977). Machiavellianism and detecting deception in facial nonverbal communication. *Towson State University Journal of Psychology, 1,* 25–31.

Bessarabova, E. (2014). The effects of culture and situational features on in-group favoritism manifested as deception. *International Journal of Intercultural Relations, 39,* 9–21. https://doi.org/10.1016/j.ijintrel.2013.09.001.

Blair, J. P., Levine, T. R., & Shaw, A. S. (2010). Content in context improves deception detection accuracy. *Human Communication Research, 36,* 423–442. https://doi.org/10.1111/j.1468-2958.2010.01382.x.

Blau, S. (2017). Indexing gay identities in American sign language. *Sign Language Studies, 18,* 5–40. https://doi.org/10.1353/sls.2017.0019.

Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14,* 189–204.

Bond, C. F., Jr., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin, 26,* 385–395. https://doi.org/10.1177/0146167200265010.

Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self-representations. *Journal of Personality and Social Psychology, 71,* 83–93. https://doi.org/10.1037/0022-3514.71.1.83.

Camden, C., Motley, M. T., & Wilson, A. (1984). White lies in interpersonal communication: A taxonomy and preliminary investigation of social motivations. *Western Journal of Speech Communication, 48,* 309–325. https://doi.org/10.1080/10570318409374167.

Cargile, A., Giles, H., & Clément, R. (1995). Language, conflict, and ethnolinguistic identity theory. *Language & Communication, 17,* 195–217.

Castillo, P. A. (2015). The detection of in cross-cultural contexts. In A. Awasthi & M. K. Mandal (Eds.), *Understanding facial expressions in communication* (pp. 243–263). New Delhi: Springer India. https://doi.org/10.1007/978-81-322-1934-7.

Cheng, K. H. W., & Broadhurst, R. (2005). The detection of deception: The effects of first and second language on lie detection ability. *Psychiatry, Psychology and Law, 12,* 107–118. https://doi.org/10.1375/pplt.2005.12.1.107.

Choi, H. J., Park, H. S., & Oh, J. Y. (2011). Cultural differences in how individuals explain their lying and truth-telling tendencies. *International Journal of Intercultural Relations, 35,* 749–766. https://doi.org/10.1016/j.ijintrel.2011.08.001.

Clementson, D. E. (2017). Truth bias and partisan bias in political deception detection. *Journal of Language and Social Psychology, 37*(4), 407–430. https://doi.org/10.1177/0261927X17744004.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995. https://doi.org/10.1037/0022-3514.70.5.979.

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivational impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12,* 177–202. https://doi.org/10.1007/BF00987487.

Ding, X. P., Wellman, H. M., Wang, Y., Fu, G., & Lee, K. (2015). Theory-of-mind training causes honest young children to lie. *Psychological Science, 26,* 1812–1821. https://doi.org/10.1177/0956797615604628.

Duh, S., Paik, J. H., Miller, P. H., Gluck, S. C., Li, H., & Himelfarb, I. (2016). Theory of mind and executive function in Chinese preschool children. *Developmental Psychology, 52,* 582–591.

Dunbar, N. E. (2018). Deception in intergroup communication. In H. Giles & J. Harwood (Eds.), *Oxford encyclopedia of intergroup communication.* New York, NY: Oxford University Press.

Dunbar, N. E., Gangi, K., Coveleski, S., Adams, A., Bernhold, Q., & Giles, H. (2016). When is it acceptable to lie? Interpersonal and intergroup perspectives on deception. *Communication Studies, 67,* 129–146. https://doi.org/10.1080/10510974.2016.1146911.

Dunbar, N. E., Jensen, M. L., Harvell-Bowman, L. A., Kelley, K. M., & Burgoon, J. K. (2017). The viability of using rapid judgments as a method of deception detection. *Communication Methods and Measures, 11,* 121–136. https://doi.org/10.1080/19312458.2016.1276893.

Ekman, P. (1997). Lying and deception. In N. Stein, P. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp. 333–347). New York, NY: Psychology Press.

Feeley, T. H., & Young, M. J. (2000). Self-reported cues about deceptive and truthful communication: The effects of cognitive capacity and communicator veracity. *Communication Quarterly, 48,* 101–119. https://doi.org/10.1080/01463370009385585.

Fitzgerald, B., Gabai, D., Jonze, S., & Smith, C. (Producers), & Smith, C. (Director). (2017). *Jim & Andy: The Great Beyond* [Motion Picture]. New York: Vice Films.

Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology, 38,* 127–150. https://doi.org/10.1111/pops.12394.

Fortman, J., & Giles, H. (2006). Communicating culture. In J. Baldwin, S. Faulknere, M. Hecht, & S. Lindsley (Eds.), *Redefining culture: Perspectives* (pp. 91–103). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Frank, M. G., Feeley, T. H., Paolantonio, N., & Servoss, T. J. (2004). Individual and small group accuracy in judging truthful and deceptive communication.

Group Decision and Negotiation, 13, 45–59. https://doi.org/10.1023/B:-GRUP.0000011945.85141.af.

Fu, G., Xu, F., Cameron, C. A., Heyman, G., & Lee, K. (2007). Cross-cultural differences in children's choices, categorizations, and evaluations of truths and lies. *Developmental Psychology, 43,* 278–293. http://psycnet.apa.org/doi/10.1037/0012-1649.43.2.278.

Furner, C. P., & George, J. F. (2014). Cultural determinants of media choice for deception. *Computers in Human Behavior, 28,* 1427–1438. https://doi.org/10.1177/0266666914543959.

Gelfand, M. J., Nishii, L. H., & Raver, J. L. (2006). On the nature and importance of cultural tightness-looseness. *Journal of Applied Psychology, 91*, 1225–1244. http://psycnet.apa.org/doi/10.1037/0021-9010.91.6.1225.

Goffman, E. (1955). On face-work: An analysis of ritual elements in social interaction. *Psychiatry, 18*, 213–231.

Goffman, E. (1967). *Interaction ritual: Essays in face-to-face behavior.* Chicago: Aldine.

Graham, J., Meindl, P., Beall, E., Johnson, K. M., & Zhang, L. (2016). Cultural differences in moral judgment and behavior, across and within societies. *Current Opinion in Psychology, 8,* 125–130. https://doi.org/10.1016/j.copsyc.2015.09.007.

Griffin, D. J., & Frank, M. G. (2018). Intercultural communication schemas of deaf and hearing adults: Visuospatial decision-making during deception detection. *Journal of Intercultural Communication Research, 47*(6), 545–563. https://doi.org/10.1080/17475759.2018.1507918.

Gudykunst, W. B., Matsumoto, Y., Ting-Toomey, S., Nishida, T., Kim, K., & Heyman, S. (1996). The influence of cultural individualism-collectivism, self construals, and individual values on communication styles across cultures. *Human Communication Research, 22,* 510–543. https://doi.org/10.1111/j.1468-2958.1996.tb00377.x.

Halevy, R., Shalvi, S., & Verschuere, B. (2014). Being honest about dishonesty: Correlating self-reports and actual lying. *Human Communication Research, 40,* 54–72. https://doi.org/10.1111/hcre.12019.

Hall, B. J., Covarrubias, P. O., & Kirschbaum, K. A. (2002). *Among cultures: The challenge of communication* (2nd ed., pp. 2–7). New York, NY: Harcourt College.

Hample, D. (1980). Purposes and effects of lying. *Southern Speech Communication Journal, 46,* 33–47. https://doi.org/10.1080/10417948009372474.

Hancock, J. T., Thom-Santelli, J., & Ritchie, T. (2004). *Deception and design: The impact of communication technology on lying behavior.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 129–134. https://doi.org/10.1145/985692.985709.

Henrich, J. (2015). Culture and social behavior. *Current Opinion in Behavioral Sciences, 3,* 84–89. https://doi.org/10.1016/j.cobeha.2015.02.001.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53,* 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109.

Hofstede, G. (1984). Cultural dimensions in management and planning. *Asia Pacific Journal of Management, 1,* 81–99. https://doi.org/10.1146/annurev.psych.53.100901.135109.

Holliday, A. (2016). Revisiting intercultural competence: Small culture formation on the go through threads of experience. *International Journal of Bias, Identity and Diversities in Education, 1,* 1–14. https://doi.org/10.4018/IJBIDE.2016070101.

Holm, H. J., & Kawagoe, T. (2010). Face-to-face lying—An experimental study in Sweden and Japan. *Journal of Economic Psychology, 31,* 310–321. https://doi.org/10.1016/j.joep.2010.01.001.

Hurley, C. M., Griffin, D. J., & Stefanone, M. A. (2014). Who told you that? Uncovering the source of believed cues to deception. *International Journal of Psychological Studies, 6,* 19–32. https://doi.org/10.5539/ijps.v6n1p19.

Kim, M., Kam, K. Y., Sharkey, W. F., & Singelis, T. M. (2008). "Deception: Moral transgression or social necessity?": Cultural-relativity of deception motivations and perceptions of deceptive communication. *Journal of International and Intercultural Communication, 1,* 23–50. https://doi.org/10.1080/17513050701621228.

Klockars, C. B. (1984). Blue lies and police placebos: The moralities of police lying. *American Behavioral Scientist, 27,* 529–544. http://dx.doi.org/10.1177/000276484027004007.

Knapp, M. L., & Comadena, M. E. (1979). Telling it like it isn't: A review of theory and research on deceptive communications. *Human Communication Research, 5,* 270–285. https://doi.org/10.1111/j.1468-2958.1979.tb00640.x.

Knapp, M. L., McGlone, M. S., Griffin, D. J., & Earnest, W. (2016). *Lying and Deception in Human Interaction* (2nd ed.). Dubuque, Iowa: Kendall Hunt.

Lapinski, M. K., & Levine, T. R. (2000). Culture and information manipulation theory: The effects of self-construal and locus of benefit on information manipulation. *Communication Studies, 51,* 55–73. https://doi.org/10.1080/10510970009388509.

Lee, K. (2013). Little liars: Development of verbal deception in children. *Child Development Perspectives, 7,* 91–96. https://doi.org/10.1111/cdep.12023.

Lee, K., Cameron, C. A., Xu, F., And, G. F., & Board, J. (1997). Chinese and Canadian children's evaluations of lying and truth telling: Similarities and differences in the context of pro- and antisocial behaviors. *Child Development, 68,* 924–934. https://doi.org/10.1111/j.1467-8624.1997.tb01971.x.

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392. https://doi.org/10.1177/0261927X14535916.

Levine, T. R. (2017a). Ecological validity and deception detection research design. *Communication Methods and Measures, 12,* 45–54. https://doi.org/10.1080/19312458.2017.1411471.

Levine, T. R. (2017b). Strong diagnostic questioning effects on deception detection accuracy with US customs agents. *Communication Research Reports, 34,* 84–87. https://doi.org/10.1080/08824096.2016.1240075.

Levine, T. R., Ali, M. V., Dean, M., Abdulla, R. A., & Garcia-Ruano, K. (2016). Toward a pan-cultural typology of deception motives. *Journal of Intercultural Communication Research, 45,* 1–12. https://www.tandfonline.com/doi/abs/10.1080/17475759.2015.1137079.

Levine, T. R., Anders, L. N., Banas, J., Baum, K. L., Endo, K., Hu, A. D. S., & Wong, N. C. H. (2000). Norms, expectations, and deception: A norm violation

model of veracity judgments. *Communication Monographs, 67*(2), 123–137. https://doi.org/10.1080/03637750009376500.

Levine, T. R., Blair, J. P., & Carpenter, C. J. (2017). A critical look at meta-analytic evidence for the cognitive approach to lie detection: A re-examination of Vrij, Fisher, and Blank. *Legal and Criminological Psychology, 23*, 7–19. https://doi.org/10.1111/lcrp.12115.

Lewis, C. C., & George, J. F. (2008). Cross-cultural deception in social networking sites and face-to-face communication. *Computers in Human Behavior, 24*, 2945–2964. https://doi.org/10.1016/j.chb.2008.05.002.

Lippard, P. V. (1988). "Ask me no questions, I'll tell you no lies": Situational exigencies for interpersonal deception. *Western Journal of Communication, 52*, 91–103. https://doi.org/10.1080/10570318809389627.

Marett, K., George, J. F., Lewis, C. C., Gupta, M., & Giordano, G. (2017). Beware the dark side: Cultural preferences for lying online. *Computers in Human Behavior, 75*, 834–844. https://doi.org/10.1016/j.chb.2017.06.021.

Markowitz, D. M., & Hancock, J. T. (2016). Linguistic obfuscation in fraudulent science. *Journal of Language and Social Psychology, 35*, 435–445. https://doi.org/10.1177/0261927X15614605.

McCaskill, C., Lucas, C., Bayley, R., & Hill, J. (2011). *The hidden treasure of black ASL: Its history and structure*. Washington, DC: Gallaudet University Press.

McCornack, S. (2015). Deceptive message production. In C. Berger, M. Roloff, S. Wilson, J. Dillard, J. Caughlin, & D. Solomon (Eds.), *The international encyclopedia of interpersonal communication* (pp. 1–5). Hoboken, NJ: Wiley Blackwell.

McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology, 33*(4), 348–377. https://doi.org/10.1177/0261927X14534656.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6*, 159–179. https://doi.org/10.1177/026540758900600202.

Nomeland, M. M., & Nomeland, R. E. (2016). *The Deaf Community in America: History in the making*. Jefferson, NC: McFarland & Company.

Oetzel, J. G., & Ting-Toomey, S. (2003). Face concerns in interpersonal conflict: A cross-cultural empirical test of the face negotiation theory. *Communication Research, 30*(6), 599–624. https://doi.org/10.1177/0093650203257841.

Park, H. S., & Ahn, J. Y. (2007). Cultural differences in judgment of truthful and deceptive messages. *Western Journal of Communication, 71*, 294–315. https://doi.org/10.1080/10570310701672877.

Park, H. S., & Levine, T. R. (2017). The effects of truth–lie base-rates on deception detection accuracy in Korea. *Asian Journal of Communication, 27*, 554–562. https://doi.org/10.1080/01292986.2017.1334074.

Seiter, J. S., Bruschke, J., & Bai, C. (2002). The acceptability of deception as a function of perceivers' culture, deceiver's intention, and deceiver-deceived relationship. *Western Journal of Communication, 66*, 158–180. https://doi.org/10.1080/10570310209374731.

Serota, K. B., & Levine, T. R. (2015). A few prolific liars: Variation in the prevalence of lying. *Journal of Language and Social Psychology, 34*, 138–157. https://doi.org/10.1177/0261927X14528804.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36*, 2–25. https://doi.org/10.1111/j.1468-2958.2009.01366.x.

Stephenson, G. R. (1967). Cultural acquisition of a specific learned response among rhesus monkeys. In D. Starek, R. Schneider, & H. J. Kuhn (Eds.), *Progress in Primatology* (pp. 279–288). Stuttgart: Gustav Fischer Verlag.

Talwar, V., Crossman, A., & Wyman, J. (2017). The role of executive functioning and theory of mind in children's lies for another and for themselves. *Early Childhood Research Quarterly, 41,* 126–135. https://doi.org/10.1016/j.ecresq.2017.07.003.

Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development, 26*, 436–444. https://doi.org/10.1080/01650250143000373.

Taylor, P. J., Larner, S., Conchie, S. M., & Menacere, T. (2017). Culture moderates changes in linguistic self-presentation and detail provision when deceiving others. *Royal Society Open Science, 4,* 1–11. https://doi.org/10.1098/rsos.170128.

ten Brinke, L., Porter, S., Korva, N., Fowler, K., Lilienfeld, S. O., & Patrick, C. J. (2017). An examination of the communication styles associated with psychopathy and their influence on observer impressions. *Journal of Nonverbal Behavior, 41,* 269–287. https://doi.org/10.1007/s10919-017-0252-5.

The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74. https://doi.org/10.1177/0022022105282295.

Ting-Toomey, S., Oetzel, J. G., & Yee-Jung, K. (2001). Self-construal types and conflict management styles. *Communication Reports, 14*, 87–104. http://psycnet.apa.org/doi/10.1080/08934210109367741.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*, 506–520. http://psycnet.apa.org/doi/10.1037/0033-295X.96.3.506.

Triandis, H. C. (1995). *New directions in social psychology. Individualism & collectivism.* Boulder, CO: Westview Press.

Triandis, H. C. (2004). The many dimensions of culture. *The Academy of Management Perspectives, 18,* 88–93. https://doi.org/10.5465/ame.2004.12689599.

Triandis, H. C. (2011). Culture and self-deception: A theoretical perspective. *Social Behavior and Personality: An International Journal, 39,* 3–14. https://doi.org/10.2224/sbp.2011.39.1.3.

Triandis, H. C., & Suh, E. M. (2002). Cultural influences on personality. *Annual Review of Psychology, 53,* 133–160. https://doi.org/10.1146/annurev.psych.53.100901.135200.

Uleman, J. S., Rhee, E., Bardoliwalla, N., Semin, G., & Toyama, M. (2000). The relational self: Closeness to ingroups depends on who they are, culture, and the type of closeness. *Asian Journal of Social Psychology, 3*, 1–17. http://psycnet.apa.org/doi/10.1111/1467-839X.00052.

Moral Dimensions of Deceptive Communication

Elaine E. Englehardt and Michael S. Pritchard

Many sentient species engage in deceptive behavior of one sort or another. The extent, if any, to which such behavior in non-human species can reasonably be subjected to moral appraisal is a matter of some disagreement. But such appraisal is commonplace in the case of humans. This chapter explores the moral dimensions of human deception.

We begin with some reflections on eighteenth-century Scottish philosopher Thomas Reid's account of moral development (Reid, 1764, pp. 196–197). That account holds that among young children's earliest social dispositions are the spontaneous tendencies openly to speak one's mind and to accept as true whatever one is told. He calls these, respectively, the *principle of veracity* and the *principle of credulity*. These dispositions are not acquired by children through experience or reasoning. Rather, they are inborn and enable children to learn from others. Neither disposition can ensure the actual truth of what is said or heard, but both are operative as the child begins to communicate with others. Although not themselves subject to moral appraisal, these dispositions are fundamental in learning language, and they help set the stage for moral appraisal.

Despite the early presence of these dispositions, children soon learn to question their reliability. They observe that not only do others often mistakenly say what is false but also they sometimes do so on purpose; that is,

E. E. Englehardt (✉)
Utah Valley University, Orem, UT, USA
e-mail: elainee@uvu.edu

M. S. Pritchard
Western Michigan University, Kalamazoo, MI, USA
e-mail: michael.pritchard@wmich.edu

they lie. Realizing this, children learn to engage in such deliberately deceptive communication themselves, sometimes in self-defense, sometimes to manipulate others, and sometimes simply to amuse themselves and others through seemingly light-hearted deception. In serious cases, learning that one has been lied to or deliberately deceived typically arouses strong negative feelings. If we engage in deliberative deception ourselves, this can give rise to strong feelings of negative self-appraisal.

Concern about the possibility of deliberately deceptive communication raises fundamental questions about *trust*. In her now classic *Lying: Moral Choice Public and Private Life*, Sissela Bok (1999) says:

> If there is no confidence in the truthfulness of others, is there any way to assess their fairness, their intentions to help or to harm? How, then, can they be trusted? Whatever matters to human beings, trust is the atmosphere in which it thrives. (p. 31)

So, when it comes to deliberately deceptive communication, the stakes can be very high. Successfully navigating one's way through these murky waters is, Bok shows, fraught with moral challenges. Wanting to maintain our credibility in the eyes of others, to avoid or minimize trouble, to evade responsibility, to gain advantage over others, and so on, are familiar and tempting gambits for the would-be liar. But, as Bok forcefully argues, we tend to overestimate the likelihood of success of such gambits, and we tend to underestimate the availability of preferable alternatives. So, we remain in the snare of such deceptive ploys.

Reid's Scottish contemporary, David Hume, dramatically poses a challenge to adhering to principles of honesty by asking his readers to consider what he calls a *sensible knave*: (Hume, 1751):

> [A]ccording to the imperfect way in which human affairs are conducted, a sensible knave, in particular incidents, may think, that an act of iniquity or infidelity will make a considerable addition to his fortune, without causing any considerable breach in the social union and confederacy. That *honesty is the best policy*, may be a good general rule; but is liable to many exceptions: And he, it may, perhaps, be thought, conducts himself with most wisdom, who observes the general rule, and takes advantage of all the exceptions. (p. 155)

Although (successful) sensible knaves might *appear* to be trustworthy, they are not. The desire not to be caught serves as a partial check on sensible knaves; after all, they are sensible, not careless. But, as long as they are as Hume describes, prepared to take advantage of opportunities to get away with wrongdoing, sensible knaves are not trustworthy. And, as Hume notes, there will be opportunities.

How would those who are not themselves sensible knaves (at least not through-and-through) respond to the threat of such knaves? Hume says of such a person:

> If his heart rebel not against such pernicious maxims, if he feel no reluctance to the thoughts of villany or baseness, he has indeed lost a considerable motive to virtue; and we may expect, that his practice will be answerable to his speculation. But in all ingenous natures, the antipathy to treachery and roguery is too strong to be counterbalanced by any views of profit or pecuniary advantage. Inward peace of mind, consciousness of integrity, a satisfactory review of our own conduct; these are circumstances very requisite to happiness, and will be cherished and cultivated by every honest man, who feels the importance of them. (pp. 155–156)

However, even those who share this response are likely to find themselves at least occasionally tempted to act like a sensible knave would; and they may give into this temptation at times. In short, honesty is not always the winner even among those who are, for the most part, honest persons. Still, explicit acknowledgment of such lapses from honesty is typically accompanied with some vestige of guilt or shame. This may decrease the chances that similar lapses will occur in the future. However, as Sissela Bok points out, in order to avoid getting caught, one might find that further deception is necessary. Or, in order to avoid self-censure, one might resort to a common form of self-denial that Adam Smith, another contemporary of Reid and Hume, calls self-deceit. According to Smith (1790, p. 117), we not only want to win the moral approval of others, but also want to see ourselves as being *worthy* of that approval. Unfortunately, this may require us to hide our moral flaws from others and from ourselves as well.

These shortcomings raise special concerns in today's world of ever-increasing dependency on the professions. Ethicist William F. May (1988, p. 408) comments on the contrast between the knowledge and expertise possessed by professionals and the relative ignorance and dependency of those of us (everyone, including professionals themselves) who must place their trust in professionals.[1] In our ignorance of that specialized knowledge and expertise, and of how it works in particular instances, we must hope that our reliance on the trustworthiness of professionals is warranted.

As May puts it, there is a "knowledge explosion" that has come with our increasing reliance on the expertise of professionals, but this is largely confined to the experts. As a consequence, it is accompanied by an "ignorance explosion" for those who do not share it, and none of us has expertise enough to be exempt from this. So, May concludes: "[Professionals] had better be virtuous. Few may be in a position to discredit [them].… [I]f knowledge is power, then ignorance is powerlessness" (p. 408). He adds: "One test of character and virtue is what a person does when no one is watching.

A society that rests on expertise needs more people who can pass that test"
(p. 408).

To illustrate the importance of professionals passing May's test, con-
sider the recent Volkswagen emissions scandal. A key Volkswagen engineer,
James Robert Laing, pled guilty to a US District Court grand jury's charge
of conspiracy to defraud the US government, to commit wire fraud, and
to violate the Clean Air Act (Chambers, Wayland, Burden, & Snell, 2016).
Laing admitted to being involved in VW's secret efforts to develop and
use a "defeat device" that enabled more than 500,000 of its vehicles from
2009 to 2015 to appear to pass US laboratory emissions tests, while they
emitted up to 40 times more nitrogen oxide than legally allowed under
ordinary driving conditions. Although not everyone who purchased one of
the VW vehicles in question did so because they thought this would be
an environmentally friendly thing to do, this was a promotional feature of
VW's advertising, and it is likely that many customers responded positively
to this.

Details about who is responsible for VW's deceptive ploy are still being
sought. However, as engineer Laing's admission makes clear, at least some
engineers share responsibility. Virtually all the codes of ethics of professional
engineering societies contain prohibitions of such dishonesty. The Preamble
of the National Society of Professional Engineers (NSPE) provides an ethical
framework for understanding why:

> Engineering is an important and learned profession. As members of this pro-
> fession, engineers are expected to exhibit the highest standards of honesty
> and integrity. Engineering has a direct and vital impact on the quality of life
> for all people. Accordingly, the services provided by engineers require honesty,
> impartiality, fairness, and equity, and must be dedicated to the protection of the
> public health, safety, and welfare. Engineers must perform under a standard of
> professional behavior that requires adherence to the highest principles of ethical
> conduct. (NSPE, 2007)

Following this preamble is a list of "Fundamental Canons" that engineers are
expected to live up to in the fulfillment of their professional duties. The first
requirement is that engineers "[h]old paramount the safety, health, and wel-
fare of the public." Another is that they issue public statements "only in an
objective and truthful manner." And yet another says, simply: "Avoid decep-
tive acts." It seems clear that James Robert Laing and any other VW engi-
neers who might have been complicit in perpetrating the emissions fraud fell
short of these three requirements and thereby also failed to satisfy the final
"Fundamental Canon," which insists that engineers are to act "honorably,
responsibly, ethically, and lawfully so as to enhance the honor, reputation, and
usefulness of the profession."

These NSPE requirements are fairly standard for professional engineering
societies. They see non-deceptive communication as a crucial responsibility

for engineers. In regard to professional responsibility, the VW emissions case makes clear that, because of their expertise and the trust invested in them, engineers have an upper hand, in many respects, in certain crucial areas of the health, safety, and welfare of the public. Of course, something similar can be said of many other professional areas on which we depend, such as medicine, pharmacy, accountancy, and law.

May does not tell his readers what passing his test requires, but it is clear that his expectation is that honesty needs to be a *virtue* if professionals are to be trustworthy. This does not necessarily require professionals always to be honest in their work, but exceptions are rare and call for ethical justification. The fact that no one may be watching when a professional employs his or her skills does not provide such a justification. Fortunately, professionals who have the virtue of honesty agree.

Despite all the attention given to honesty and truthfulness in most codes of ethics, the subtlety and nuances of these and closely related notions typically are left to the reader to determine. Further, there is no detailed discussion in the codes of why dishonesty in professional practice is regarded to be generally unacceptable. We will now explore some of these matters further. We will discuss five basic forms of dishonesty: (1) outright lying; (2) deliberate deception that, strictly speaking, does not involve lying; (3) withholding or hiding information; (4) secretive and illicit use of confidential information; and (5) the failure to seek out the truth when this is one's responsibility. Although dishonesty is displayed in each of these five forms, not all of them involve direct action. However, they all raise questions about ethics in communication.

Five Forms of Dishonesty

1. Outright Lying

Dishonesty is readily associated with lying, although it is not restricted to this. It is not difficult to come up with clear-cut examples of lying that are familiar to all of us (e.g., the child who is caught with his hand in the cookie jar, with telling crumbs around his mouth, but who denies that he has gotten into the cookies). However, defining lying is not easy. One reason is that saying something that is not true is not necessarily a lie, not even in cases where one should "know better." If a child incorrectly says that $3+7 = 9$, this may simply be a mistake, not a lie. If an engineer mistakenly conveys incorrect test results on soil samples, he or she is not lying even if what is reported is not true. To lie in such cases, one must intentionally, or at least knowingly, try to convey what one believes to be false or misleading information. But even here complications arise. Someone may offer information that he or she believes to be false, even though it is actually true (e.g., one may say that $7 \times 8 = 56$, while believing that the true answer is 54). If we know that this is what has

happened, we still may fault him or her for making the mistake, but we may still wonder if we should call this a lie.

Complicating matters further, making false statements is not the only way of deceiving others. Gesturing and nodding, as well as making misleading, but true statements, can create false impressions in conversation, even if no outright lie has been made. Despite these complications, an outright lie typically uses words to convey something the liar believes is false or seriously misleading, and this is done with the intention to deceive others.

If successful, a lie may lead those who are taken in by it to do things that they would not otherwise have done. This, we can say, is a failure to respect the deceived person's capacity to make informed decisions on his or her own. It might also result in unjustified harm to the deceived, or to third parties to whom the deception is passed by those who are originally deceived.

2. Deliberate Deception that, Strictly Speaking, Does Not Involve Lying

Suppose that Rick tries to impress his employer or customer by discussing technical matters in a way that creates the false impression that he has expertise that he knows he lacks. This is an act of deliberate deception, even if Rick is not outright lying. Rick can also deliberately misrepresent the value of certain products or designs by praising their advantages inordinately while not mentioning their shortcomings. It might be thought that this form of deception is less objectionable from an ethical point of view than outright lying. However, this may not be so. Those deceived may still complain that their capacity to decide for themselves was not respected. They may still suffer serious harms, and third parties can suffer as well.

The media is depended on to play an important role in helping the public avoid being victimized by deception. However, insofar as it might be influenced by funding from private donors, foundations, businesses, and political organizations, the media can itself be a powerful source of such deception. Another major source of media support is advertising. One of the more famous TV advertisements years ago was for Wrigley's Doublemint chewing gum, which charmed audiences with a clever jingle sung by smiling twins. Even today the "double your pleasure, double your fun" tune is recalled by many. Whatever the words of this jingle might literally have meant, having them delivered by actual twins seemed to be important to the Wrigley company.

In contrast, today's media technology easily enables TV to create ads featuring such well-known celebrities as NBA basketball star Chris Paul sharing adventures with his fabricated "twin brother." Do viewers actually believe that they are seeing twins interacting in support of the product being promoted? Many might say that it doesn't matter if there are actual twins or not. No one should actually be taken in by this false representation; it might be said. If some are, that is their fault for being so naïve. In the "good old days"

of the Doublemint twins, it might have mattered to viewers, but not today. How is this to be explained? Both the old and new ads are entertaining, and everything portrayed is seemingly harmless. Is this enough to justify the fabrication? Without attempting to settle this issue, it does seem fair to say that some sort of justification of the deceptive ad should be available. Some ads, for example, those that provide a subscript saying that the people featured in them are not actors, suggest that truthfulness in advertising does still count for something. However, saying that they are not actors is not the same as saying that they are not simply pretending that they like the product.

The Chris Paul ad illustrates how easily technology can "manufacture" false appearances. Such ads may seem innocuous enough. However, charges that the media cannot be trusted because it now specializes in "fake news" (also technologically feasible) can shake the confidence the public traditionally has had in the "Fourth Estate" as a reliable source of information that can enable citizens to be knowledgeable about matters of practical importance. This is no trivial matter.

3. Withholding or Hiding Information

Omitting or withholding information can be another type of deceptive behavior. If Jane deliberately withholds information from her superior that she knows would shed negative light on a project she is promoting to her superior, she engages in serious deception even if she is not lying. For this, she may be blamed as fully as if she had told an outright lie.

But one might have good intentions in resorting to this sort of deception. Author James Herriot is well known for his stories about his adventures as an English veterinarian. In *All Creatures Great and Small*, he recounts a story about visiting Mrs. Tompkins, a frail, elderly woman whose budgie, Peter, had become uncharacteristically inactive (Harriot, 1973, pp. 298–302). Thinking that clipping the budgie's beak might help, Harriot reached into Peter's cage and gently picked him up. However, Peter fell limp in his hand and died from heart failure. Harriot realized that, due to her poor hearing and eyesight, Mrs. Tompkins had not noticed what happened to Peter. He also worried that telling Mrs. Tompkins what had just happened might be devastating to her. So, he quickly decided it would be better to rush Peter out of her house, saying that he could do the clipping better in his office.

Harriot quietly disposed of Peter and found a look-alike substitute budgie at a bird store. He then placed the new budgie in Peter's cage and assured Mrs. Tompkins that now all would be well with Peter. It took Harriot several months to summon up the courage to revisit Mrs. Tompkins and check on how she and the new bird were faring. He feared that she may have discovered this Peter was an imposter. Mrs. Tompkins excitedly greeted Harriot at the door: "You know, you wouldn't believe it, he's a different bird." Harriot swallowed hard and responded: "Is that so? In what way?" "Well," she

replied, "He's so active now. Lively as can be. You know he chatters to me all day long. It's wonderful what cutting a beak can do."

This, we might say, is a case of *benevolent* deception. Our first response might be to admire Harriot for his inventiveness in a difficult situation. However, further reflection might raise some doubts about whether Harriot took the best course of action. Could Mrs. Tompkins have been preparing herself for the worst with Peter? Was she expecting his death soon? Was it important to her that she have the opportunity to say good-bye to him? Perhaps this is not such a simple, heartwarming, and amusing story after all. In any case, there may be more to it morally than (at least initially) meets the eye.

4. Secretive and Illicit Use of Confidential Information

Respecting confidentiality requires *not* revealing information that others may want, but have no legitimate right to demand. The protection of such information is a key to many successful business ventures, particularly in the area of competitive bidding. Consider this fictional case. Company X has prepared a competitive bid for a major government contract. Gwen has been given the task of hand delivering the confidential bid. In doing this, Gwen finds herself seated on an airplane next to another businessperson. They strike up a conversation and soon discover that they have much in common. It turns out that Gwen and her seatmate are in identical professions with similar educational backgrounds. Some time into the flight, Gwen's seatmate confesses that he despises the firm that employs him and wonders if there are any openings at Gwen's place of employment. She explains to him that she works for a small, fledgling company, and that she is hand delivering a bid that could make or break the future of the company. In fact, if it were awarded the bid, there would most likely some job openings. Surprised, the seatmate explains that he, too, is delivering a bid for this same contract. In fact, he says, "I'd better take a few minutes to prepare a bit more for my presentation." He excuses himself, opens his travel case, takes out what is obviously his company's bid, and begins to read it. A few minutes later, he excuses himself to go to the restroom and stretch for a few minutes. He leaves the bid on his seat and goes to the back of the plane.

Gwen now realizes she has an unexpected opportunity to look at the competing bid. It seems obvious to her that this is just what her seatmate is hoping she will do. Why else, she wonders, would he make his company's bid so accessible? Is he hoping that this might create a job opportunity for him should Gwen's company win the bid? If Gwen does examine the bid, would she then be tempted to alter her company's bid? Gwen decides that looking at the bid would violate the principles of fair play that are integral to the sort of competitive bidding process to which her company has committed itself. As for her seatmate's apparent desire to join her company's firm, Gwen concludes that he is not someone to trust. If he is willing to harm his company

by giving Gwen access to the bid, what else might he be willing to do? He is not someone that she would want as a co-worker.

5. Failure to Seek Out the Truth When This is One's Responsibility

An honest researcher is committed to seeking and finding the truth, even when it is not certain what the outcome of further research might reveal. Suppose a biochemist suspects that some of the data received from the test lab might be inaccurate or seriously misleading. Using the results without inquiring further into their accuracy seems deceptive, even if it does not involve outright lying.

    As already noted, it should not be assumed that outright lying is always more serious than other forms of deliberate deception. In some cases, much more harm may actually result from other forms of deception than from outright lying. Our discussion of different forms of dishonesty has focused primarily on the degree to which truth is being actively distorted or misused rather than the seriousness of the consequences of those actions.

HONESTY AND CANDOR

May and Hume see suitably constrained honesty as a virtue. As such, it should be distinguished from candor. A moral requirement of total candor would tolerate people being tactless in telling one another what they really think about their opinions, their looks, their choices of food, their habits, and so on. Some tact and even reticence are marks of politeness and civility. Sacrificing this in the name of promoting "truth" would seem to be a heavy price to pay in interpersonal relations. In professional life, a requirement never to conceal truth would mean that engineers, physicians, lawyers, and other professionals should not protect confidentiality or proprietary information. Doctors could never misrepresent the truth to their patients, even when there is strong evidence that this is what the patients prefer and that the truth could be devastating.

DISHONESTY AND VIRTUE

Despite the harms that complete candor could bring with it, lying and the various other ways of being dishonest are generally wrong. But, as a virtue, honesty does require good judgment. Most would agree that lying to a would-be murderer regarding the whereabouts of a would-be victim is justified and does not by itself render one a dishonest person.

    George Washington is often lauded for his honesty. The famous story of young George admitting to his father that he chopped down his favorite cherry tree is commonly interpreted as supporting the view that he would

never lie. However, his line, "I cannot tell a lie, father" could be understood contextually as "I cannot tell a lie about the wrong I have done you, father." The story concludes, "To the end of his life [George Washington] was just as brave and honorable as he was that day as a little boy" (Esenwin & Stockard, 1993, p. 606). This suggests that his honesty in this case supports the view that he was always brave and honorable, not that he never lied. Chopping down the cherry tree was an act he needed to own up to as his. Lying about it would be dishonorable. That is not the same as saying that lying is always dishonorable. He could still hold that he would be justified in lying in order to save his father from being murdered. This, his father could acknowledge, would be a courageous and honorable act, too—one of which he could rightly be proud.

Obviously, not all instances of dishonesty involve stakes as high as whether to save an innocent person from being murdered. Nor does resorting to dishonesty in less extreme circumstances necessarily mean that one is acting like a sensible knave. However, it can put one's virtue of honesty on the line. Consider the case of Norm Lewis, who in 1968 was a 51-year-old doctoral candidate in history at the University of Washington (Rivera, 2000). While taking his final exam in the program, he excused himself to go to the restroom. Apparently fearing that he would fail if he did not take a secret look at his notes, he did look at them. It is not known whether his fear was well grounded. He might well have passed even without cheating. After all, in subsequent years he published several books on different aspects of religion. Perhaps he simply had a moment of panic during the final exam. Taking a few deep breaths might have restored his confidence.

However, for the next 32 years, Lewis told no one about what he had done, not even his spouse. At age 83, he decided to confess, and he wrote to the president of the university, admitting that he had cheated and that he had regretted it ever since.

Commenting on the case, Jeanne Wilson, president of the Center for Academic Integrity, remarked, "I think there is an important lesson here for students about the costs of cheating. He has felt guilty all these years, and has felt burdened by this secret, believing that he never really earned the degree he was awarded." Wilson concluded that the University of Washington should not take action against Lewis, given his confession, his age, and that fact that, after all, he did complete his coursework and a dissertation.

Wilson did not say that Norm Lewis was morally justified in what he did. It was, she agreed, wrong for him to have looked at his notes; but she thought nothing would be gained from depriving him of his degree at this late date, or from taking any other action against him. As for Lewis's self-assessment, apparently he felt his moral integrity was still on the line. Not to own up to others his wrongdoing would be, he may have thought, as dishonorable as George Washington lying to his father about the cutting down the cherry tree. Admittedly, the confession came rather late, but it was an attempt to make amends as best he could while he still had some time to do so.

Dishonesty and Respect for Persons

One fundamental moral concern about dishonesty is its failure to show respect for the ability (and right) of those deceived to make informed decisions about matters of significance to them. As moral agents, we value ourselves as being to some extent capable of formulating and pursuing goals and purposes of our own. That is, we see ourselves as somewhat autonomous in the moral sphere. This is a matter of considerable importance to us in our everyday lives, as members of families, as friends or lovers, as customers shopping for goods, or even in casual exchanges with those we do not know well.

It is also of considerable importance to us insofar as we interact with or depend on professionals for our well-being. This is well illustrated in medical practice insofar as physicians take care to ensure that their patients give informed consent to the treatments they will receive. This involves efforts to help patients understand the likely consequences of this or that course of action, both in the short run and in relation to their longer-term life plans. It also involves trying to help patients avoid making decisions that are primarily the result of undue coercive factors such as stress, illness, or family pressures. Finally, physicians are expected to familiarize patients with information about different available options for treatment and their expected consequences. This can pose special challenges. To be informed appropriately, patients must not only *have* the relevant information but also *understand* it. This can present special challenges to those whose expertise is not widely understood by others, and whether these challenges are well handled can be subject to controversy.

Dishonesty from a Utilitarian Perspective

A utilitarian perspective commends actions that tend to promote human happiness and well-being and avoid the opposite. Making precise determinations of this sort can be very difficult, especially if one must look far into the future in order to predict likely consequences. For example, good engineering can contribute to utilitarian goals by providing designs for the creation of reliable buildings, bridges, electronic devices, automobiles, and many other things on which our society depends. However, even good engineering cannot guarantee that unfortunate results will be avoided. There are storms, earthquakes, terrorist attacks, and other disasters that obviously can interfere with accomplishing intended results.

Even welcome technological innovations, such as cell phones and other communication devices, can have unwelcome, unintended consequences (e.g., car accidents caused by the distraction of the device). The difficulty of reliably predicting unwelcome consequences of technological innovations highlights one of the major challenges facing utilitarian thinking. However, not fully accepting the responsibility of seriously trying to make reliable predictions invites even worse consequences. At least this much can be said in

behalf of the utilitarian concern for maximizing good consequences and minimizing bad ones.

Dishonesty in Research

The most striking advances in science, engineering, and communication have depended on good research. Furthermore, successful research is seldom the result of the work of solitary individuals. It depends directly or indirectly on the reliable work of researchers who pass the results of their work on to future researchers, as well as on the reliable work of teams of current researchers whose members cooperate internally with one another and who, largely through publications, share the results of their work with other researchers. For this to work well, basic honesty in research is essential. If researchers falsify data or omit crucial data, then others, including other researchers, cannot reliably depend on their results. This can undermine the relations of trust on which a community of research is founded. Just as a designer who is untruthful about the strength of materials specified for a building fails to protect the building from harm, a researcher who falsifies the data reported in a professional journal threatens harm to those who rely directly or indirectly on the reported results.

This harm does not stop with other researchers. Dishonest research can also undermine informed decision making by business executives, government officials, policymakers, and citizens generally who depend on the knowledge and judgment provided by researchers in making decisions. If the reports made available to them are unreliable, then their ability to make good recommendations and decisions about such matters is compromised. Insofar as the unreliability of these reports is due to dishonesty in research, moral (if not legal) culpability is involved, and public welfare may be put at risk.

There are several well-known types of dishonesty in research. These are labeled "research misconduct" by federal agencies that fund research, such as the National Science Foundation (NSF) and the National Institutes of Health (NIH). Colleges and universities that sponsor research funded by such agencies have policies with provisions that prohibit such misconduct. Falsification of data, fabrication of data, plagiarism, and inappropriate attributions of authorship are among the most common forms of research misconduct.

Falsification of data involves distorting data by smoothing out irregularities or presenting only those data which fit one's favored theory and discarding the rest. A well-known dispute about whether results were falsified pivots around Robert A. Millikan's oil drop experiment early in the twentieth century (Holton, 1978, pp. 161–224). Millikan received a Nobel Prize for his published research that was credited with demonstrating the uniformity of the charge of electrons. Later, however, it was discovered that, despite claiming that his findings were based on all 189 trials he ran, his notebooks indicate that his report did not include data on 49 of his trials. Millikan defenders

point out that sometimes equipment does not function as it should and that, in those circumstances, omission of data may be appropriate. Critics reply that such problems should be indicated in published reports. Further experimentation confirmed Millikan's conclusions, but some still question whether Millikan's own experiment should be cited as showing this.

Fabrication of data involves inventing data and even reporting the results of experiments that were never conducted. The case of psychologist Stephen Breuning illustrates how far the harm from doing this can reach (Sprague, 1998, p. 33).[2] In December 1983, Dr. Robert Sprague of the University of Illinois wrote a lengthy letter, accompanied with an even lengthier set of appendices, to the National Institute of Mental Health (NIMH) that documented the fraudulent work submitted to him by his young associate researcher, Steven Breuning. Although Sprague and Breuning were collaborators, they did not conduct their research in the same place. As head of research on this NIMH funded project, Sprague conducted his research primarily at the University of Illinois at Urbana-Champaign. Breuning claimed he was reporting on research he had conducted in a Coldwater, Michigan mental facility on the effects psychotropic medication has on children in need of help. However, Sprague claimed, Breuning simply made up the data.

Three months after Sprague sent his letter to NIMH, Breuning admitted to fabricating his data. Nevertheless, it took more than five years for the case finally to be settled. Ironically, Sprague found himself to be the first one questioned in the lengthy investigation that followed his making charges against Breuning. His regularly funded NIMH research was severely curtailed, he was subjected to threats of lawsuits, and he had to testify before a US House of Representatives committee. In the midst of this investigation, Sprague's wife died after a lengthy bout with diabetes. In fact, Sprague said, his wife's serious illness was one of the major factors prompting him to report Breuning to NIMH. Concerned about how dependent his diabetic wife was on reliable research and medication, Sprague was also sensitive to the dependency that children, and vulnerable populations in general, have on the trustworthiness of those who do research in areas pertaining to their well-being.

Reflecting further on Breuning's fabricated research, Sprague expressed some concerns about other possible victims—namely, other psychologists and researchers who had collaborated with Breuning without being aware that he had fabricated data. Western Michigan University psychologist Alan Poling, who at one time had Breuning as an M.A. student, has written about the consequences of Breuning's misconduct for others with whom he collaborated in research (Poling, 1992, pp. 140–157). Strikingly, Poling points out that during a critical time of research, Breuning was a contributor to more than one-third of all published research on the psychopharmacological areas in which he specialized. Of course, it does not follow that all of Breuning's publications were based on fabricated data, but determining which were and which were not would be a time-consuming, daunting task. Meanwhile, those who,

in fact, relied on Breuning's work in support of their own research had reason to be concerned.

Especially since researchers are dependent on one another to advance reliable research of their own, honesty among researchers is a needed virtue. It is essential for, not only the mutual respect that is needed among researchers, but also for those who, although not themselves members of the community of researchers, make decisions based on the published findings of researchers. Those decisions can have profound consequences for those far removed from the world of scientific research, such as young children with special needs and their caregivers.

FABRICATING SUPPORT

If one needs the support of others in order to gain recognition, is it all right to fabricate that support? Apparently, this was once the view of Timothy J. Cooney, author of *Telling Right from Wrong* (Cooney, 1985). Although intensely interested in philosophical questions, Cooney never attained an academic position in philosophy, nor did he acquire a Ph.D. in this subject. Convinced that publishers would not accept the writings of someone without such academic credentials, Cooney adopted a rather desperate tactic. He made up a letter filled with praise for the book, as well as a detailed critical analysis, signed the name of well-known Harvard philosopher Robert Nozick to the letter and sent it off to Random House, along with the draft of his book. Random House contracted with Cooney to publish his book. Just as the book was being set up for publication, someone from Random House happened to meet Nozick at a party and asked him to comment further on this book that he had praised so highly. Nozick replied that he had known nothing about either the book or its author.

In defending his act of deception, Cooney is quoted as saying to a *New York Times* reporter that he was actually very proud of it: "It has no moral or ethical implications…. It has nothing to do with the end of the world. I would call it 'vigorous gameplay,' and considering my book is now in galleys I'm very proud of what I did. Whether it makes book form or not, it's a lot more advanced than it would have been without the forgery" (McDowell, 1984). Apparently, Cooney thought that, according to the theory he defends in his book, such an act of deception was not a serious enough matter to count as wrongdoing. Random House disagreed and regretfully withdrew its plans to publish Cooney's book.

Cooney then took his draft to Prometheus, which agreed to publish it—on the condition that he write an afterword in which he would discuss his fraudulent act in relation to his moral theory. In light of the highly publicized critical responses to his initial attempt to win the attention of Random House, Cooney reassessed what he had done. Although his behavior was not of the sort that, even if widely practiced, would threaten the survival of the world, or even some significant portion of it, its negative consequences for others would be quite unwelcome.

In preparing his final draft for Random House, Cooney received extensive editorial help from Jonathan Lieberson, then an assistant professor of philosophy at Barnard College. After learning that Cooney had fabricated the Nozick letter and that the book was rejected by Random House, Lieberson commented: "What is tragic here is that in some sense this is a victimless crime, or a crime in which Cooney is the only victim…. His book is a good one and his argument, while controversial, can stand on its own. But there still is a real issue of violation of trust, and that is quite appalling" (McDowell, 1984). Indeed, the issue of trust is crucial, and this is not just about Cooney. Had the fraud not been exposed prior to going to press, Random House would have published the book. Nozick's "endorsement" would have enhanced its sales, and many would falsely have believed that Nozick thought highly of its contents. Not only would Nozick have been the victim of fraud, his own philosophical views might well have been seriously misrepresented in the letter. For any philosopher, let alone one as celebrated as Robert Nozick, this is no small matter. Finally, to excuse Cooney's behavior is to excuse similar behavior on the part of anyone in circumstances such as his. The reach of this generalization is far beyond the world of Timothy Cooney, Robert Nozick, and others in the rather limited circle of those who might have been affected by what Cooney attempted to get away with. Although this is not an issue that puts the survival of the world at stake, it is a morally serious one.

Plagiarism

Another kind of concern about deception in regard to writing, research, and publication is *plagiarism*, the adoption of the work of others without acknowledgment of the appropriate authorship. This amounts to a kind of theft of the work of others by taking false credit for that work. This need not involve lifting an entire book or article written by someone else. It can be as little as a single sentence or a small gathering of words. In any case, it is a deceptive act of communication. Plagiarism may more commonly occur in classroom writing assignments, but it also occurs beyond academic settings in the world of publication. If plagiarized work is cited on one's CV, this is yet another form of deception, as is listing fake activities or accomplishments in general.

Concluding Thoughts

In the course of discussing human dimensions of deception, we hope to have highlighted those areas of communication that illuminate how morality comes into play. In ordinary interpersonal relations, the media, the world of business, professional life, and policy-making, honesty in communication is key to trustworthiness. This may still leave room for some justifiable deception. However, betrayal, disloyalty, lack of respect and care for persons, and the failure to safeguard the public and its welfare are all disvalues that can come from deception in human communication. So, the topic of deception in the human sphere is filled with moral significance.

Notes

1. Much of what follows in this article is based on the main lines of our "Trust and Reliability," Chapter 5 of *Engineering Ethics: Concepts and Cases*, 6th ed., a book co-authored with C. E. Harris, Raymond James, and the late Michael Rabins (Cengage, 2017, pp. 97–120). We (Englehardt and Pritchard) had primary responsibility for authoring this chapter.
2. For an account of this case, see Robert L. Sprague, "The Voice of Experience," *Science and Engineering Ethics*, Vol. 4, 1, 1998, pp. 33–44. The discussion of this case here is based on our analysis in Chapter 5 of *Engineering Ethics: Concepts and Cases*, 6th ed., co-authored with C. E. Harris, Raymond James, and the late Michael Rabins (Cengage, 2017).

References

Bok, S. (1999). *Lying: Moral choice in public and private life* (2nd ed., p. 31). New York: Vintage Books.

Chambers, J., Wayland, M., Burden, M., & Snell, R. (2016, September 9). VW engineer's guilty pleas could signal more indictments. *Detroit News*. https://www.detroitnews.com/story/business/autos/foreign/2016/09/09/vw-charges/90118226/. Accessed April 9, 2018.

Cooney, T. (1985). *Telling right from wrong*. Buffalo: Prometheus.

Esenwin, J. B., & Stockard, M. (1993). George Washington and the cherry tree. In W. Bennett (Ed.), *The book of virtues* (p. 606). New York: Simon & Schuster.

Harriot, J. (1973). *All creatures great and small* (pp. 298–302). New York: Bantam.

Harris, C. E., Pritchard, M. S., James, R. W., Englehardt, E., & Rabins, M. J. (2017). *Engineering ethics: Concepts and cases* (6th ed.). Boston: Cengage.

Holton, G. (1978). Subelectrons, presuppositions, and the Millikan-Ehrenhaaft dispute. *Historical Studies in the Physical Sciences, 9*, 161–224.

Hume, D. (1751). *An enquiry concerning the principles of moral understanding* (T. L. Beauchamp, Ed., p. 155). New York: Oxford University Press, 1998.

May, W. F. (1988). Professional virtue and self-regulation. In J. Callahan (Ed.), *Ethical issues in professional life* (p. 408). Oxford: Oxford University Press.

McDowell, E. (1984, September 13). Book on morality poses dilemma. *New York Times*.

NSPE. (2007). *National society for professional ethics*. NSPE Code of Ethics. https://www.nspe.org/sites/default/files/resources/pdfs/Ethics/CodeofEthics/Code.

Poling, A. (1992). The consequences of fraud. In D. Miller & M. Hersen (Eds.), *Research fraud in the behavioral and biomedical sciences* (pp. 140–157). New York: Wiley.

Reid, T. (1764). *An inquiry into the human mind*, Part VI, Sect. 24 (R. Beanblossom & K. Lehrer, Eds., pp. 196–197). Indianapolis: Hackett.

Rivera, R. 2000. 32 years after cheating, missionary confesses. *The Seattle Times*. Accessed April 8, 2018. http://community.seattletimes.nwsource.com/archive/?date=20000724&slug=4033407.

Smith, A. (1790). *The theory of moral sentiments* (D. D. Raphael & A. L. MacFie, Eds.). Indianapolis: Liberty Press, 1984 (p. 117).

Sprague, R. L. (1998). The voice of experience. *Science and Engineering Ethics, 4*(1), 33–44.

Effects of Deception on the Deceiver: An Interdisciplinary View

Richard Reardon, Annette L. Folwell,
Jode Keehr and Trevor Kauer

Deception research touches many critical areas that are key to understanding social interactions such as adaptation, cognition and emotion, self-concept, and effective communication. Research on deception has focused on who lies, why people lie, why they are sometimes successful at lying, factors contributing to the detection of deception, and the immediate interpersonal fallout for those deceived (i.e., the targets). Relatively less attention has been paid to the impacts of deception on the deceiver. There are notable exceptions to this claim, such as Buller and Burgoon's (1996) interpersonal deception theory (IDT) where impacts on the deceiver account for one of the three components of the theory, along with the target and observers of the deception. There are other bodies of research that address impacts on deceivers, but this work is often oblique to other issues that are considered primary. We believe that a gathering of theory and research regarding impacts on the deceiver

R. Reardon (✉) · A. L. Folwell · J. Keehr
University of Idaho, Moscow, ID, USA
e-mail: rreardon@uidaho.edu

A. L. Folwell
e-mail: folwell@uidaho.edu

J. Keehr
e-mail: jkeehr@uidaho.edu

T. Kauer
Texas State University, San Marcos, TX, USA
e-mail: tbk17@txstate.edu

with elaboration could be useful. In this chapter, we address explicitly how the act of engaging in deception can have direct effects on the deceiver and we begin in a place that is not often a starting point for considering deception—the brain.

The chapter begins with a brief overview of some foundational concepts and distinctions in deception research. We discuss how deception may have emerged as an adaptive mechanism and offer an overview of relevant neural processes in which the first effects on the deceiver occur. Then, the ways cognitive and emotional loads impact the deceiver will be considered, and the relationship between deception and self-understanding will be discussed. Subsequently, we propose ways that the deceiver's understanding of the reality of the communicative moment might be affected. Lastly, we conclude by linking our view to current communication research focusing on some immediate effects of deception on the deceiver.

FOUNDATIONS

To establish a foundation for this chapter, basic understandings regarding the deception process must be recognized including the prevalence of deception, a working definition of deception, truth bias, relationship closeness, and deception tactics. First, deception is an everyday, commonplace occurrence and not an unusual event (DePaulo & Kashy, 1998). Although deception is an everyday occurrence, we recognize the presence of the truth bias: Honesty is valued more, and the majority of people are honest most of the time (e.g., Buller & Burgoon, 1996). Further, this chapter will concentrate on "consequential" deceptions, which conceptually overlap with Walczyk, Harris, Duck, and Mulay's (2014) notion of "serious lies" but is more inclusive. A "serious lie" has a significant impact (e.g., endangerment or betrayal) on the target (Walczyk et al., 2014); while we include those deceptions, we widen our view to include any deception that has a meaningful consequence for the deceiver as well as the target. Both consequential deceptions and benign deceptions, where interpersonal consequences are minimal, draw on the same adaptive, neural, and interpersonal mechanisms. Lies seem to be more common in high stakes relationships (e.g., Hample, 1980; Millar & Tesser, 1988), but occur in all kinds of relationships, including those when the stakes are low (e.g., DePaulo & Kashy, 1998; Kalbfleisch, 2001; Williams, 2001).

Deception has both verbal and nonverbal components, but the nonverbal aspects of interpersonal deception have garnered more research attention than verbal aspects (Burgoon & Qin, 2006). For instance, Burgoon, Proudfoot, Scheutzler, and Wilson (2014) found that "deceptive nonverbal communication patterns are shorter, less complex, more redundant, and less diverse than exhibited by truthful nonverbal patterns" (p. 349). More recently, there has been an increasing interest in verbal components of deception examining lexical, syntactic, and meta-content features by comparing truthful versus

deceptive discourse; deceivers' messages become longer, more specific, more diverse, more "active," simpler, less redundant, and less immediate over time than truth messages (Burgoon & Qin, 2006). Also, we acknowledge that deception is accomplished through numerous communication channels and systems that may impact the deceiver; for example, a deceiver in a computer-mediated medium is more dominant, involved, relaxed, and active when compared to a deceiver in a face-to-face situation (Dunbar et al., 2013).

The number of times a specific deceptive act occurs may affect the deceiver as well. Burgoon (2015) speculates:

> ….as the number and duration of utterances related to an issue increases, the more cognitively challenging it should be to lie, inasmuch as one must remember what has been said previously, create consistency among utterances, reconcile what is being said with a potentially growing population of known facts, make decisions about which truthful details to divulge, decide what kinds of deception to enact, whether to change strategies (e.g., from concealment to equivocation), and so forth. (p. 6)

Therefore, the possible repetitive nature of deception may be a factor that impacts the effects on the deceiver, but additional research is needed to determine the consequences of repetition.

While there are many specific motivations for deception, they can be organized into two archetypes: Deceivers can either use a reticent strategy, which appears to be passive and commonly known as "lying by omission," or they may use a more active strategy that appears more manipulative (Burgoon et al., 2014). Indeed, numerous studies examined how deception can be used for self-serving purposes such as to save face (e.g., Horan & Booth-Butterfield, 2013), avoid conflict (e.g., Metts, 1989; Peterson, 1996), and manage impressions (e.g., Peterson, 1996; Seiter, Bruschke, & Bai, 2002). Other researchers have focused on the role competent deception plays in being an effective and polite communicator (e.g., Dunbar et al., 2016), managing relational boundaries (e.g., Derlega, Metts, Petronio, & Margulis, 1993), and repressing negative thoughts to maintain a sense of autonomy (Cole, 2001).

While we cannot discuss all types of motivations for invoking deception in this chapter, we recognize that there are many different motivations for deception (Ekman, 2009; Lindskold & Walters, 1983; Peterson, 1996; Seiter et al., 2002). Just as deception can be employed as an explicit or implicit tactic, it can be used with good or evil intent or as a deflective act. Regardless of motivation, any discussion that attempts to explain the impact of deception on the deceiver must begin with the notion that deception involves choices; in IDT, deceptions are *intentional* (Buller & Burgoon, 1996). While some scholars have made a point of addressing "self-serving" lies as part of a repertoire of tools available to serve the self (Burgoon, 2015; Whitty & Carville, 2008), the adaptive viewpoint is that all deceptions are self-serving.

Importantly, the counterpart of deception is honesty. As noted earlier, there is a truth bias that serves the deceiver and the target. In the next section, we discuss the adaptive significance of deception and truth-telling.

BIOLOGY OF DECEPTION

Adaptive Significance of Deception

Our brains are very much like those of our primate relatives, but with one significant difference: Humans have a higher neocortex volume, relative to the rest of the brain, than other primates (Dunbar, 2007). This greater volume is associated with a number of advantages, including computation, logic, and reasoning, but especially social skills such as coalition formation and complexity, play, mating behaviors, and power relationships (Dunbar, 2007). Importantly, according to Byrne and Corp (2004), one of the adaptive social gains is the ability to "tactically deceive," that is, the deceiver draws from their normal repertoire behaviors needed to produce a misinterpretation by the target that advantages the deceiver. Thus, as humans evolved so have social skills, including the ability to deceive.

Krebs and Denton (1997) and von Hippel and Trivers (2011) offer further perspective on the adaptive service that deception provides. Competent deception aids in gaining and protecting resources, increasing social power, maintaining relationships, and enhancing one's sense of self. Deception also buffers the self-concept from threats of detection; when deceivers employ deceptive tactics, they simultaneously self-deceive becoming less aware of the deception influenced by the strong pressure to be honest. The underlying mechanism is a desire to maintain a positive self-view while deceiving the target. Self-deception is necessary to be a competent deceiver, whereas truth-telling makes no demands on the deceiver (Baumeister, 1993). Self-deception helps the deceiver avoid self-betrayal in the sense that it is similar to other dissociative processes such as repression, suppression, and hypnosis. As discussed later in the chapter, the brain is organized as an assemblage of multiple systems, some of which can control and influence others (von Hippel & Trivers, 2011; note that Buller & Burgoon, 1996, use "dissociation" in a slightly different, but compatible sense).

Researchers have suggested deception and the ability to detect deception evolved simultaneously (Krebs & Denton, 1997; von Hippel & Trivers, 2011). As reflected in the truth bias, society values truth over deception so the ability to detect deception is a useful adaptation. Like many adaptations, imperfection is assumed for both deception and detection; it is best not to be perfect at deceiving others nor at detecting deception. Failure to detect allows the deceiver to keep the self-deception intact and the social connection to continue (Dunbar, 2007). The advancement of collaborative social interests is more important than complete deception success or exceptional

detection accuracy. Deception evolved as a tool to be used in certain situations, whereas detection of deception evolved because it is useful in determining what is real. Also, there are constraints on deception; one restriction is the ability of a target to detect the deception, and another is the truth bias. The truth bias has important adaptive advantages; if truth was not the dominant complement to deception, then a person's world would be unpredictable and unstable. The truth requires less cognitive effort (Spence, 2004) and is associated with fewer long-term health problems (ten Brinke, Lee, & Carney, 2015); thus, societal and cultural norms of honesty may have biological underpinnings.

Neuroscience of Deception

There are a number of tools available to help pinpoint the role various parts of the brain play in active social interaction. The ancients understood that trauma or illness affecting specific brain areas was informative; they recognized that behavioral, emotional, and cognitive changes could be correlated with areas of damage. In the modern era, analysis of physiological changes revealed by polygraph recordings (e.g., heart rate, blood pressure, galvanic skin resistance, and respiration) suggests certain changes in brain activity. Even routine blood testing reveals influences on emotion that are tied to brain physiology. However, recent astonishing breakthroughs have been made possible by advances in imaging and scanning. In particular, functional magnetic resonance imaging (fMRI) reveals changes in neural activity *over time*. These devices and methods permit a number of conclusions and informed inferences regarding the neuroscience of deception.

We have assumed that deception is an adaptive tool; thus, there should be brain areas developed for deceptive action. While it would be convenient to think of the brain as a single grand system, we now believe it to be multiple systems in which activities are executively controlled and awareness resides in narratives constructed as the systems interact to ensure survival (Gazzaniga, 2000, 2011; Lieberman, 2013; von Hippel & Trivers, 2011). There is no single, special deception brain area; rather, there are multiple areas that act in concert to accomplish deception. These same areas, when tasked differently, are also associated with truth-telling, self-examination, social comparison, and other social cognitions and acts.

Explicitly or implicitly, the deceiver comes to an understanding of their role in the communicative moment. In the service of self, the deceiver must process and encode inputs, retrieve selected past information, and then weigh the short-term and long-term costs of either being truthful or engaging in deception. In doing so, the deceiver considers their sense of self and weighs the interpersonal risks of engaging in deception. These activities are not without the influence of social and cultural norms about what is right and wrong.

The self-concept that resides in an individual's brain is a reflection of societal norms and values acquired through socialization (Lieberman, 2013).

At the point of employing truth or deception, the presence of truth bias suggests that the default action will be truthfulness. There will be times, however, when the deceiver has an explicit or implicit sense that a deception might serve them, and perhaps serve the target as well (Spence et al., 2004). In a specific moment, a person chooses to be either a deceiver or a truth-teller. A deceiver considers what the target may find plausible and then weaves a deception to suit the deceiver's goals. Even the simplest deception must be crafted, and this generative activity employs multiple areas of the brain (Frith & Frith, 2006; Lieberman, 2013). The deception is communicated and the target responds. The same systems involved in crafting the original deception are simultaneously monitoring the target's reaction and creating a response. Thus, the cognitive load on the deceiver becomes intense.

As self-deception is a function of the dissociation of truth and deception in memory, one of the ways the brain accomplishes self-deception is by the simple act of inhibition (Levy & Anderson, 2012). While the brain permits explicit awareness of the deception, it makes the truth less available to awareness (Levy & Anderson, 2012; von Hippel & Trivers, 2011). Gazzaniga (2011) proposed that there is a left-hemispheric "interpreter" that supports the creation of meaningful explanations of information gathered from multiple brain areas. The evidence that a narrative explanation is created is compelling. Importantly, the narratives created are only as complete as the information that is readily available. If unwanted memories, such as those that reflect the truth, have been suppressed, they are less likely to be included in any narrative.

Self-deception is compounded by the very malleability of our memories. Memory is susceptible to distortion (Loftus, 1991). Memories can be altered by type of lie told, and a person may even begin to believe a falsehood if the truth has been denied multiple times (Vieira & Lane, 2013).

We have discussed what the brain does at a functional level. We turn now to pinpoint brain areas that change in activity level when processes involved in deception occur: self-examination, cognitive ability/intelligence, creativity, other perspective taking, cognitive inhibition and suppression, monitoring accuracy of self and others' cognitions, and instantiating an emotional reaction. Those areas are identified in Fig. 6.1.

The frontal lobes of the brain are responsible for cognitive executive activity. In the course of deception, the areas of the frontal lobes most often associated are the prefrontal cortex (PFC) and the anterior cingulate cortex (ACC). As shown in Fig. 6.1, the PFC sits in the front of the brain and the ACC surrounds the corpus callosum, which joins the two hemispheres.

The PFC, ACC, and portions of the parietal lobes are reliably associated with self-knowing and self-reflection (Fiske & Taylor, 2017); these regions
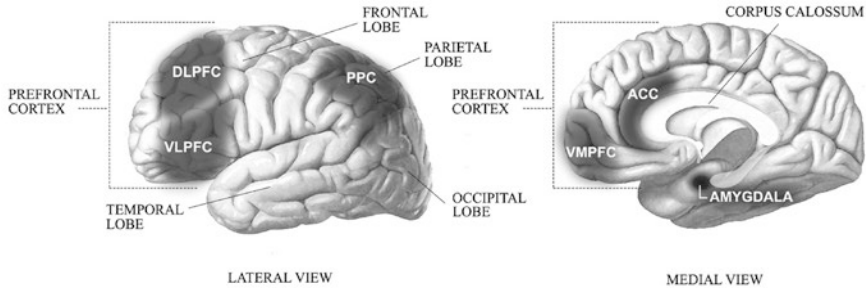
Fig. 6.1 Left: lateral view of the human brain denoting locations of the dorsolateral prefrontal cortex (DLPFC), post-parietal cortex (PPC), ventrolateral prefrontal cortex (VLPFC). Right: medial view of the human brain denoting locations of the ventromedial prefrontal cortex (VMPFC), the anterior cingulate cortex (ACC), and the amygdala

show greater activity when we think about ourselves and others. The lateral (i.e., outer) portions of the PFC are associated with reasoning, working memory, and social cognition. Toward the rear of the PFC, the dorsolateral prefrontal cortex (DLPFC) is involved in generating novel responses and behaviors (Frith, Friston, Liddle, & Frackowiak, 1991; Lieberman, 2013; Spence et al., 2004). The DLPFC and the posterior parietal cortex (PPC) are selectively involved with working memory (Christ, Van Essen, Watson, Brubaker, & McDermott, 2009). Increased activation in the ventrolateral prefrontal cortex (VLPFC) is related to response inhibition (Christ et al., 2009; Starkstein & Robinson, 1997), rehearsed lies (Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003), and suppression of truthful responses (Spence et al., 2001). Both the DLPFC and VLPFC are active in suppressive processes inhibiting memory encoding and retrieval, and have been linked to "motivated forgetting" (Anderson & Hanslmayr, 2014).

The ACC serves both cognitive and emotional functions. In addition to its involvement in self-knowing, it helps regulate conflicting sensory input and helps identify erroneous information (e.g., Bush, Luu, & Posner, 2000; Phan, Wager, Taylor, & Liberzon, 2002). In doing so, it monitors anticipated positive or negative outcomes (Apps & Ramnani, 2014). Consequently, there is increased activity in the ACC during deception (Abe et al., 2006; Ganis et al., 2003; Langleben et al., 2002).

The medial prefrontal cortex (MPFC), located in the PFC toward the central fissure, is key to sustaining a sense of self by monitoring the varying intensity levels of activity from other systems and connecting them to emotional experience. Lieberman (2013) suggests that this "consultation" of the MPFC, in combination with connections to the emotional system, is where moral judgment enters the deception process.

Many areas of the brain are associated with mood (Phan et al., 2002), though the limbic system is generally acknowledged as the key structure for emotional processing. Specifically, the amygdala is involved with rapid assessment of emotional consequences, monitoring salient cues from the environment delivered by the sensory systems. Signals from the amygdala are dispatched to the MPFC, which regulates the emotional response (Lindquist, Wager, Kober, Bliess-Moreau, & Barrett, 2012). Fundamentally, arousal is produced. Potential harm, fear of detection, and violation of truth bias affect the labeling of that arousal, and an emotional state consistent with the label is felt. Recently, Garrett, Lazzaro, Ariely, and Sharot (2016) identified the biological underpinnings of what is known as "the slippery slope." Signals from the amygdala weaken with each self-serving deception; over time, lies become less distressing to the deceiver and deceptions are easier to employ.

COGNITIVE AND EMOTIONAL LOADS

The deceiver must craft the deception, mentally dissociate from the truth, control both verbal and nonverbal behaviors that might generate suspicion from the target, and think about what the target currently knows and might learn from future interactions. These activities create cognitive load, and that load affects how well the deceiver adjusts to the target's reaction and processes immediate changes in the context and target's direct response. Increased cognitive load can also narrow the focus of attention, influence planning and anticipation, and affect and be affected by the processing of the complexity of the broader context (Buller & Burgoon, 1996; Sporer, 2016; von Hippel & Trivers, 2011; Walczyk et al., 2014). Sporer (2016), citing Wegner's (1994) Ironic Processing Theory, adds that attempts to dissociate truth from deception in memory may make the truth memory stronger and more likely to intrude into working memory.

There is also an increase in emotional load when a person engages in deception. A deceiver must manage changes in arousal and the attributed emotional states—primarily fear, guilt, and shame. Additionally, Mather and Sutherland (2011) address selective effects of arousal, finding that arousal narrows the focus of attention. Especially if the arousal has uncertain causes, it compels interpretation (Mather, Clewett, Sakaki, & Harley, 2015). The narrow focus and pressing need to attribute an emotional state can affect the construction and presentation of the deception, adding to processing load.

Positively labeled emotional states (e.g., happiness and joy) tend to favor heuristic, schematic, top-down processing, whereas negatively labeled states (e.g., fear, guilt, or shame) lead to greater attentional focus on the situational information. However, the increased attention to situational information does not guarantee accuracy in processing. Ambady and Gray (2002) state the advantages of being more attuned to situational information may be offset by the inability to interpret target and bystander cues. The impact of negatively

valanced arousal may be complex; the deceiver may focus too narrowly on social and situational cues, and be biased by the emotional state so much that the ability to call on useful schematic and heuristic processing is reduced. In a deceptive moment, the deceiver may find that they are less able to formulate the next argument, and less able to understand the target's reply, or interpret the target's other verbal or nonverbal cues (Mather & Sutherland, 2011; Walczyk et al, 2014).

Forgas (1995, 2003) suggests that emotional states intrude into the cognitive system, directly impacting cognitive construction. Deceptive statements are constructions, and emotional content influences their construction; Forgas (1995) calls this "infusion." In other words, the experienced emotion primes, or colors, the construction of the statement; for example, feelings of fear or guilt would infuse the cognitive constructions with content that is congruent with these emotions. Forgas (1995) found that infusion is more likely as constructive activity increases. Having engaged in deception, the deceiver feels fear, guilt or shame; that emotional state is manifested in the construction of follow-up verbal responses and nonverbal behaviors, and interpretation of the target's and bystanders' language and behaviors.

Self-perception

Protagoras (ca. 490–420BC) claimed that each of us is the measure of our own reality; simply put, we use our self as a yardstick to make decisions about relationships, undertake tasks, reconcile our past, and make future plans. This view is generally held today (Fiske & Taylor, 2017; Forgas, Williams, & Wheeler, 2003; Sedikides, Gaertner, Luke, O'Mara, & Gebauer, 2013); therefore, fundamental adaptive advantages of deception are an enhancement of the self and preservation of self-concept.

As we noted, self-deception co-occurs with deception (von Hippel & Trivers, 2011); therefore, one advantage of self-deception is that it reduces cognitive load by allowing the deceiver to be less aware of the deceptive act. This self-deception serves the deceiver by helping to avoid sending cues that might cause suspicion. If the deception was successful, it produces self-enhancement; if the deception was detected, it allows the deceiver to claim ineptitude, confusion, or ignorance.

Self-deception makes some aspects of the self-concept unavailable, which can affect the quality of the interaction. Moreover, self-deception influences other aspects of memory. It can make searches more or less selective and time constrained, impact attention and vigilance, create bias in the interpretation of incoming information and retrieval of stored memories, and affects the evaluation of the past leading to elaborate rationalization (von Hippel & Trivers, 2011).

Because the deception might occur in different interactions, the deception must be repeated in order for it to be sustained. The deceiver is then more susceptible to errors in reality monitoring, a type of memory source confusion (Johnson, 2006; Johnson & Raye, 1981; Reardon & Doyle, 1995). Perceptions, as opposed to cognitive generations, have more contextual detail as part of the memory; the memories of generations, on the other hand, are richer in memory *process* information. When a memory is activated, contextual detail may be added as the deceiver elaborates on the deception to make it more believable. To the deceiver, the deception can feel more and more like a reality when it is repeated and information is added or modified (cf. Goff & Roediger, 1998).

CHANGES IN SENSE OF REALITY

In the deceptive moment, the deceiver's brain is actively hiding or altering aspects of the deceptive content or circumstances. Furthermore, there is a concomitant impact on how information is being processed and retrieved, verbal and nonverbal behaviors are adjusted explicitly or implicitly and, if the deception situation is to be sustained, a deceiver's understanding of the truthfulness of the deception will change. Along with these considerations, the differing perspectives and motives of the deceiver, the target, and observers can create confusion (Burgoon, Buller, Floyd, & Grandpre, 1996). The deceptive moments may feel less real, which may impact the communication.

Fundamental to any action are the outcomes, and outcomes may impact a person's sense of reality. As suggested, an implicit or explicit monitoring of the self is necessary if the goal of deception is to self-enhance or protect the self-concept; then, the deception may create circumstances that make the world feel less real. This logic follows an attributional model developed by Brickman (1978) in which one's sense of reality is influenced by: (a) the correspondence between one's behavior and its outcomes; and (b) the correspondence between one's behavior and their internal standards and values. A situation feels most real when the consequences of the behavior are appropriate, *and* that behavior is consistent with internal standards and values. Typically, when deception occurs, internal standards are violated and a state of unreality is experienced that is characterized by alienation or estrangement from others. On occasion, deception can feel more like role-playing than living in the moment because the deception was designed to avoid some calamity *and* it violated internal norms and values. A deception may produce a sense of unreality significant enough to impact our perceptions of others and events at that moment. Reardon, Keehr, Folwell, and Hackworth (2018) found partial support for these notions; in their study, introduction to deception led participants to a sense of unreality that was manifested as suspicion about others' truthfulness.

Interpersonal Deception

While we contend that all deception has a self-serving purpose, we acknowledge there may be other goals present. Deceptions can serve the interest of the target of the deception or the relationship itself but, as Knapp (2006) noted, a deceiver will participate in deceptive communication because they have an interest (i.e., purpose) in the conversation. Typically, deceivers will tell more self-centered than other-oriented lies with the exception of those pertaining to feelings, which are described as other-oriented to benefit the receiver (DePaulo, Kashy, Kirkenfol, Wyer, & Epstein, 1996). Metts (1989) found that a common motivation for deception in romantic relationships, when compared to friendships, is out of concern for either the partner or relationship. The motivation for telling lies is at the heart of determining which lies are more acceptable. Deceptive messages told for altruistic reasons are perceived as more acceptable than those told for selfish or malicious reasons (e.g., Dunbar et al., 2016; Seiter et al., 2002), and "white lies" are far more common and acceptable than "big lies" that may have more profound consequences (Serota & Levine, 2015). Regardless of whether the lies are perceived to be more or less acceptable, the process of deceiving another person has real consequences on the deceiver.

Relational context also impacts the effects of deception on the deceiver. In an organizational setting, Dunbar et al. (2014) reported an increase in deceivers' perception of their own power and cognitive demands, whereas their truthful partners reported a decrease in perceptions of their own power and cognitive demand. Individuals in a position of power may deceive their subordinates because they need to hide information, motivate subordinates, or feel they have the right to be deceptive; those in lower power positions may use deception to assert their own agency or evade aversive consequences stemming from their powerless status (Lindsey, Dunbar, & Russell, 2011). Thus, changes in perceived power of the deceiver can be a common effect of deception and this change may be dependent on their success in the deception.

The use of deceptive communication will impact the overall quality of the relationship (Kalbfleisch, 2001) because lying violates both relational (DePaulo & Kashy, 1998) and moral partner expectations (Peterson, 1996); consequently, deceivers should experience some aversive emotional state (Horan & Dillow, 2009). Some effects of deception are felt at the individual level, such as negative emotions and communication competence, while other consequences occur at the relational level, such as commitment and satisfaction. Relatedly, self-deception and the perception of self as a deceiver have a role in creating and sustaining an aversive emotional state in the deceiver. While deceivers used different message strategies (Kalbfleisch, 2001), they felt less closeness and pleasantness in their relationships (Lawson, 2000; DePaulo et al., 1996).

People who engage in deception experience negative emotions particularly when lying in high stakes relationships (DePaulo & Kashy, 1998; DePaulo et al., 1996). Given that we presume others to be honest (i.e., the truth bias), deceivers know their goal of deception and the actual deceptive message are both in conflict with social expectations. When being deceptive, the deceiver is aware of violating the social norm of being honest, which contributes to their feelings of fear (Ekman & Friesen, 1969; Zuckerman, DePaulo, & Rosenthal, 1981). When a deceiver lies, arousal is created internally and the deception elicits more distress, discomfort, and associated physiological changes including changes in heart rate, body temperature, and blood pressure (Smyth & Pennebaker, 2008).

Ekman (2009) argued that, along with fear, deception is linked to guilt and shame. Further, Ekman (2009) stated that while fear of detection emerges from the possibility of being caught in deception, guilt and shame are elicited from going against one's values and beliefs (Seiter & Brunschke, 2007). Guilt is associated with possible objections to the deceiver's act of deception (Baumeister, Stillwell, & Heathteron, 1995) and occurs regardless of whether or not there is an audience (Ekman, 2009). Guilt is associated with increased feelings of insecurity, sadness, and helplessness (Horan & Booth-Butterfield, 2011; Jones & Kugler, 1993) as well as regret, misery, resentment, and loneliness (Tangney, Wagner, Hills-Barlow, Marschall, & Gramzow, 1996). Guilt also decreases the feelings of comfort and competence in the deceiver (Horan & Booth-Butterfield, 2011; Jones & Kugler, 1993). Similarly, shame is the distressing self-examination with corresponding feelings of "shrinking and being small" (Tangney et al., 1996); shame does require the presence of others, who are potential judges of the deceiver (Ekman, 2009). Shame is positively associated with maladaptive anger and hurt responses, including physical and verbal aggression (Tangney et al., 1996). Essentially, both shame and guilt are self-conscious emotions that result from deceptive communication and appear to be complex experiences related to other important emotional responses (Andersen & Guerrero, 1998; Horan & Booth-Butterfield, 2011).

Individuals experience guilt when they deceive and may also feel shame if they consider themselves to be a deceptive person (Horan & Booth-Butterfield, 2011). Interestingly, deceivers report feelings of guilt with both imagined and actual deception; thus, guilt is associated with the thought process not necessarily the deceptive act (Gozna, Vrij, & Bull, 2001; Hample, 1980; Peterson, 1996). In addition, the amount of guilt felt is related to the severity of the deception; when people tell "white lies," instead of "blatant lies," they report less guilt (Peterson, 1996), and shame is less likely to be reported than guilt (Tangney, 1992). After deception, if the individual feels guilt more than shame, they are more likely to attempt to fix the situation to alleviate the guilty feelings; if the individual feels shame more than guilt, they are more likely to withdraw from the situation (Tangney, Wagner, & Gramzow, 1992).

Targets sometimes become suspicious of deception, and this suspicion is associated with probing questions and comments aimed at resolving doubt. Probing will provoke more arousal and may lead to changes in the deceiver's nonverbal and verbal communication behaviors. Nonverbal behaviors, such as increased eye contact, maintaining baseline levels of facial animation, and giving more signs of agreement, are executed in the hope of supporting the deception and maintaining a positive image (Buller & Burgoon, 1994; DePaulo, Stone, & Lassiter, 1985). In addition, verbal reactions to suspicion include longer responses, and more speech errors and pauses (Buller, Comstock, Aune, & Strzyzeweki, 1989). Interestingly, Arciuli, Mallard, and Villar (2010) asserted the verbal utterance of "um" may not be used to fill pauses or correct speech disfluency/error; rather, it may carry lexical status to make the deceiver seem more authentic while deceiving under increased arousal and cognitive load. Therefore, deceivers may attempt to be perceived as more truthful by increasing their involvement and creating a positive image when they are confronted by suspicion and probing messages (Buller, Strzyzewski, & Comstock, 1991). Thus, there are immediate impacts on the deceiver in the communicative setting that emerge from interaction. The deceiver's emotions evoke behaviors that can affect the target's reaction. The ability of the deceiver to maintain the deception in an extended interaction and avoid detection could also be further impaired by the target's suspicion. As the evidence shows, when a deception is sustained over extended time and it starts to feel real, the deceiver might feel some relief.

Conclusion

We have argued that deception is an adaptive tool, constrained by societal norms and values, which affects the self and self-concept by explicit or implicit awareness of possible detection of the deceptive act. Deception occurs when brain systems are employed in a concerted manner. The brain registers an incongruency between truth and deception, and attempts to suppress the truth as the deceiver simultaneously presents the alternate reality of the deception to the targets. The emotional system is activated through arousal manifesting as fear, guilt, shame, or a combination of these emotions in the deceiver; additionally, long-term health can be impacted. As the deception is generated, emotions are triggered, relevant attributions are made, and cognitive load increases. Cognitive functioning is impaired. The deceiver is unable to accurately evaluate the exchange and, in particular, the responses of the target. Accompanying self-deception, a sense of unreality can develop as the deceiver copes with the suppression of truth. The deceiver must monitor the target for signs of detection while communicating, and all of this occurs while the deceiver experiences an aversive emotional state. During the communication process, the effects of deception on the deceiver appear as changes in perception of one's own power as well as

relational closeness. Additionally, the emotions of fear, guilt, and shame can lead to suspicion and probing by the target that can further aggravate situational cognition. While much research has focused on the detection of deception, the keys to detection may be in understanding the very complex role of the deceiver.

References

Abe, N., Suzuki, M., Tsukiura, T., Mori, E., Yamaguchi, K., Itoh, M., et al. (2006). Dissociable roles of prefrontal and anterior cingulate cortices in deception. *Cerebral Cortex, 16,* 192–199. https://doi.org/10.1093/cercor/bhi097.

Ambady, N., & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin slice judgments. *Journal of Personality and Social Psychology, 83,* 947–961. https://doi.org/10.1037/0022-3514.83.4.947.

Anderson, P. A., & Guerrero, L. K. (1998). *Handbook of communication and emotion.* San Diego, CA: Academic Press.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in cognitive sciences, 18*(6), 279–292.

Apps, M. A. J., & Ramnani, N. (2014). The anterior cingulate gyrus signals the net value of others' rewards. *Journal of Neuroscience, 34,* 6190–6200. https://doi.org/10.1523/jneurosci.2701-13.2014.

Arciuli, J., Mallard, D., & Villar, G. (2010). "Um, I can tell you're lying:" Linguistic markers of deception versus truth-telling in speech. *Applied Psycholinguitics, 31,* 397–411. https://doi.org/10.1017/s0142716410000044.

Baumeister, R. F. (1993). Lying to yourself: The enigma of self-deception. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 166–183). New York: Guilford Press.

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1995). Personal narratives about guilt: Role in action control and interpersonal relationships. *Basic and Applied Social Psychology, 17,* 173–198. https://doi.org/10.1080/01973533.1995.9646138.

Brickman, P. (1978). Is it real? In J. Harvey, W. Ickes, & R. Kidd (Eds.), *New directions in attribution research* (pp. 5–34). Hillsdale, NJ: Lawrence Erlbaum.

Buller, D. B., & Burgoon, J. K. (1994). Deception: Strategic and nonstrategic communication. In J. A. Daly & J. M. Wiemann (Eds.), *Strategic interpersonal communication* (pp. 191–223). Hillsdale, NJ: Erlbaum.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x.

Buller, D. B., Comstock, J., Aune, R. K., & Strzyzewski, K. D. (1989). The effect of probing on deceivers and truthtellers. *Journal of Nonverbal Behavior, 13,* 155–170. https://doi.org/10.1007/bf00987047.

Buller, D. B., Strzyzewski, K. D., & Comstock, J. (1991). Interpersonal deception: I. Deceivers' reactions to receivers' suspicions and probing. *Communication Monographs, 58,* 1–24. https://doi.org/10.1080/03637759109376211.

Burgoon, J. K. (2015). When is deceptive message production more effortful than truth-telling? A baker's dozen of moderators. *Frontiers in Psychology, 6,* 1–6. https://doi.org/10.3389/fpsyg.2015.01965.

Burgoon, J. K., Buller, D. B., Floyd, K., & Grandpre, J. (1996). Deceptive realities. *Communication Research, 23,* 724–748. https://doi.org/10.1177/009365096023006005.

Burgoon, J. K., Proudfoot, J. G., Schuetzler, R., & Wilson, D. (2014). Patterns of nonverbal behavior associated with truth and deception: Illustrations from three experiments. *Journal of Nonverbal Behavior, 38,* 325–354. https://doi.org/10.1007/s10919-014-0181-5.

Burgoon, J. K., & Qin, T. (2006). The dynamic nature of deceptive verbal communication. *Journal of Language and Social Psychology, 25,* 76–96. https://doi.org/10.1177/0261927x05284482.

Bush, G., Luu, P., & Posner, M. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215–222. https://doi.org/10.1016/s1364-6613(00)01483-2.

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society of London, 271,* 1693–1699. https://doi.org/10.1098/rspb.2004.2780.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex, 19,* 1557–1566. https://doi.org/10.1093/cercor/bhn189.

Cole, T. (2001). Lying to the one you love: The use of deception in romantic relationships. *Journal of Social and Personal Relationships, 18,* 107–129. https://doi.org/10.1177/026540750118100.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74,* 63–79. https://doi.org/10.1037/0022-3514.74.1.63.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995. https://doi.org/10.1037/0022-3514.70.5.979.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Deceiving and detecting deceit. In B. R. Schlenker (Ed.), *The self and social life* (pp. 323–370). New York: McGraw-Hill.

Derlega, V. J., Metts, S., Petronio, S., & Margulis, S. T. (1993). *Self-disclosure.* Newbury Park, CA: Sage.

Dunbar, N. E., Gangi, K., Coveleski, S., Adams, A., Bernhold, Q., & Giles, H. (2016). When is it acceptable to lie? Interpersonal and intergroup perspectives on deception. *Communication Studies, 67,* 129–146. https://doi.org/10.1080/10510974.2016.1146911.

Dunbar, N. E., Jensen, M. L., Bessarabova, E., Burgoon, J. K., Bernard, D. R., Harrison, K. J., et al. (2014). Empowered by persuasive deception: The effects of power and deception on dominance, credibility, and decision making. *Communication Research, 41,* 852–876. https://doi.org/10.1177/0093650212447099.

Dunbar, N. E., Jensen, M. L., Burgoon, J. K., Kelly, K. M., Harrison, B. J. A., & Bernard, D. R. (2013). Effects of veracity, modality, and sanctioning on credibility assessment during mediated and unmediated interviews. *Communication Research, 42,* 649–674. https://doi.org/10.1177/0093650213480175.

Dunbar, R. I. M. (2007). The social brain hypothesis and its relevance to social psychology. In J. P. Forgas, M. G. Haselton, & W. von Hippel (Eds.), *Evolution and the social mind: Evolutionary psychology and social cognition* (pp. 21–31). New York: Psychology Press.

Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (4th ed.). New York: W.W. Norton.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1,* 49–98. https://doi.org/10.1515/semi.1969.1.1.49.

Fiske, S. T., & Taylor, S. E. (2017). *Social cognition: From brain to culture*. Thousand Oaks, CA: Sage.

Forgas, J. P. (1995). Mood and judgment: The affect intrusion model. *Psychological Bulletin, 117,* 39–66. https://doi.org/10.1037/0033-2909.117.1.39.

Forgas, J. P. (2003). Affect and the "social mind": Affective influences on strategic interpersonal behaviors. In J. P. Forgas, K. D. Williams, & L. Wheeler (Eds.), *The social mind: Cognitive and motivational aspects of interpersonal behavior* (pp. 46–71). New York: Cambridge University Press.

Forgas, J. P., Williams, K. D., & Wheeler, L. (2003). *The social mind: Cognitive and motivational aspects of interpersonal behavior*. New York: Cambridge University Press.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50,* 531–534. https://doi.org/10.1016/j.neuron.2006.05.001.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. J. (1991). A PET study of word finding. *Neuropsychologia, 29,* 1137–1148. https://doi.org/10.1016/0028-3932(91)90029-8.

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex, 13,* 830–836. https://doi.org/10.1093/cercor/13.8.830.

Garrett, N., Lazzaro, S. C., Ariely, D., & Sharot, T. (2016). The brain adapts to dishonesty. *Nature Neuroscience, 19,* 1727–1732. https://doi.org/10.1038/nn.4426.

Gazzaniga, M. S. (2000). Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain, 123,* 1293–1326. https://doi.org/10.1093/brain/123.7.1293.

Gazzaniga, M. (2011). *Who's in charge? Free will and the science of the brain*. New York: Harper.

Goff, L. M., & Roediger, H. L., III. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition*, *26,* 20–33.

Gozna, L. F., Vrij, A., & Bull, R. (2001). The impact of individual differences on perceptions of lying in everyday life and in a high stake situation. *Personality and Individual Differences, 31,* 1203–1216. https://doi.org/10.1016/s0191-8869(00)00219-1.

Hample, D. (1980). Purposes and effects of lying. *Southern Speech Communication Journal, 46,* 33–47. https://doi.org/10.1080/10417948009372474.

Horan, S. M., & Booth-Butterfield, M. (2011). Is it worth lying for? Physiological and emotional implications of recalling deceptive affection. *Human Communication Research, 37,* 78–106. https://doi.org/10.1111/j.1468-2958.2010.01394.x.

Horan, S. M., & Booth-Butterfield, M. (2013). Understanding the routine expression of deceptive affection in romantic relationships. *Communication Quarterly, 61,* 195–216. https://doi.org/10.1080/01463373.2012.751435.

Horan, S. M., & Dillow, M. R. (2009). Deceivers and emotion: The relationships among deceptive message type, relational qualities, and guilt and shame. *Atlantic Journal of Communication, 17,* 149–165. https://doi.org/10.1080/15456870903156126.

Johnson, M. K. (2006). Memory and reality. *American Psychologist, 61,* 760–771. https://doi.org/10.1037/0003-066x.61.8.760.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67–85. https://doi.org/10.1037/0033-295x.88.1.67.

Jones, W. H., & Kugler, K. (1993). Interpersonal correlates of the Guilt Inventory. *Journal of Personality Assessment, 61,* 246–258. https://doi.org/10.1207/s15327752jpa6102_6.

Kalbfleisch, P. J. (2001). Deceptive message intent and relational quality. *Journal of Language and Social Psychology, 20,* 214–230. https://doi.org/10.1177/0261927x01020001010.

Knapp, M. L. (2006). Lying and deception in close relationships. In A. L. Vandelisti & D. Perlman (Eds.), *The Cambridge handbook of personal relationships* (pp. 517–532). New York: Cambridge University Press.

Krebs, D. L., & Denton, K. (1997). Social illusions and self-deception: The evolution of biases in person perception. In J. A. Simpson & D. T. Kenrick (Eds.), *Evolutionary social psychology* (pp. 21–48). Mahweh, NJ: Lawrence Erlbaum.

Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., et al. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *NeuroImage, 15,* 727–732. https://doi.org/10.1006/nimg.2001.1003.

Lawson, T. J. (2000). Are kind lies better than unkind truths? Effects of perspective and closeness of relationship. *Representative Research in Social Psychology, 24,* 11–19.

Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *Journal of Neuroscience, 32,* 16785–16794. https://doi.org/10.1523/jneurosci.2640-12.2012.

Lieberman, M. D. (2013). *Social: How our brains are wired to connect.* New York, NY: Crown Publishers.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences, 35,* 121–143. https://doi.org/10.1017/s0140525x11000446.

Lindskold, S., & Walters, P. S. (1983). Categories for acceptability of lies. *The Journal of Social Psychology, 120,* 129–136. https://doi.org/10.1080/00224545.1983.9712018.

Lindsey, L. L. M., Dunbar, N. E., & Russell, J. C. (2011). Risky business or managed event? Perceptions of power and deception in the workplace. *Journal of Organizational Culture, Communications and Conflict, 15,* 55–79.

Loftus, E. F. (1991). Made in memory: Distortions in recollection after misleading information. *Psychology of Learning and Motivation—Advances in Research and Theory, 27,* 187–215. https://doi.org/10.1016/s0079-7421(08)60124-3.

Mather, M., Clewett, D., Sakaki, M., & Harley, C. (2015). Norepinephrine ignites local hot spots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences, 39,* 1–100. https://doi.org/10.1017/S0140525X15000667.

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science, 6,* 114–133. https://doi.org/10.1177/1745691611400234.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6,* 159–179. https://doi.org/10.1177/026540758900600202.

Millar, K. U., & Tesser, A. (1988). Deceptive behavior in social relationships: A consequence of violated expectations. *The Journal of Psychology, 122,* 263–273. https://doi.org/10.1080/00223980.1988.9915514.

Peterson, C. (1996). Deception in intimate relationships. *International Journal of Psychology, 31,* 279–288. https://doi.org/10.1080/002075996401034.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage, 16,* 331–348. https://doi.org/10.1006/nimg.2002.1087.

Reardon, R., & Doyle, S. (1995). The self-concept and reality judgments: Memory, memory monitoring, and internal-external correspondence. *Social Cognition, 13,* 1–24. https://doi.org/10.1521/soco.1995.13.1.1.

Reardon, R., Keehr, J., Folwell, A. L., & Hackworth, C. (2018, March). *Deception affects feelings of reality.* Paper presented at the meeting of the Southern Society for Philosophy and Psychology, San Antonio, TX.

Sedikides, C., Gaertner, L., Luke, M. A., O'Mara, E. M., & Gebauer, J. E. (2013). A three-tier hierarchy of self-potency: Individual self, relational self, collective self. *Advances in Experimental Social Psychology, 48,* 235–295. https://doi.org/10.1016/b978-0-12-407188-9.00005-3.

Seiter, J. S., & Bruschke, J. (2007). Deception and emotion: The effects of motivation, relationship type, and sex on expected feelings of guilt and shame following acts of deception in United States and Chines samples. *Communication Studies, 58,* 1–16. https://doi.org/10.1080/10510970601168624.

Seiter, J. S., Bruschke, J., & Bai, C. (2002). The acceptability of deception as a function of perceivers' culture, deceiver's intention, and deceiver-deceived relationship. *Western Journal of Communication, 66,* 158–180. https://doi.org/10.1080/10570310209374731.

Serota, K. B., & Levine, T. R. (2015). A few prolific liars: Variation in the prevalence of lying. *Journal of Language and Social Psychology, 34,* 138–157. https://doi.org/10.1177/0261927x14528804.

Smyth, J. M., & Pennebaker, J. W. (2008). Exploring the boundary conditions of expressive writing: In search of the right recipe. *British Journal of Health Psychology, 13,* 1–7. https://doi.org/10.1348/135910707x260117.

Spence, S. A. (2004). The deceptive brain. *Journal of the Royal Society of Medicine, 97,* 6–9. https://doi.org/10.1258/jrsm.97.1.6.

Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. *NeuroReport, 12,* 2849–2853. https://doi.org/10.1097/00001756-200109170-00019.

Spence, S. A., Hunter, M. D., Farrow, T. F. D., Green, R. D., Leung, D. H., Hughes, C. J., et al. (2004). A cognitive neurobiological account of deception. *Philosophical Transactions of the Royal Society of London, 359,* 1755–1762. https://doi.org/10.1098/rstb.2004.1555.

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology, 7,* 1–12. https://doi.org/10.3389/fpsyg.2016.00420.

Starkstein, S. E., & Robinson, R. G. (1997). Mechanism of disinhibition after brain lesions. *The Journal of Nervous and Mental Disease, 185,* 108–114.

Tangney, J. P. (1992). Situational detennninants of shame and guilt in young adulthood. *Personality and Social Psychology Bulletin, 18,* 199–206. https://doi.org/10.1177/0146167292182011.

Tangney, J. P., Wagner, P., & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology, 101,* 469–478. https://doi.org/10.1037/0021-843x.101.3.469.

Tangney, J. P., Wagner, P., Hills-Barlow, D., Marshall, D. E., & Gramzow, R. (1996). Relation of shame and guilt to constructive versus destructive responses to anger across the life span. *Journal of Personality and Social Psychology, 70,* 797–809. https://doi.org/10.1037/0022-3514.70.4.797.

ten Brinke, L., Lee, J. J., & Carney, D. R. (2015). The physiology of (dis)honesty: Does it impact health? *Current Opinion in Psychology, 6,* 177–182. https://doi.org/10.1016/copsyc.2015.08.004.

Viera, K. M., & Lane, S. M. (2013). How you lie affects what you remember. *Journal of Applied Research in Memory and Cognition, 2,* 173–178. https://doi.org/10.1016/j.jarmac.2013.05.005.

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34,* 1–56. https://doi.org/10.1017/s0140525x10001354.

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology, 34,* 22–36. https://doi.org/10.1016/j.newideapsych.2014.03.001.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101,* 34–52. https://doi.org/10.1111/j.1467-9280.2006.01697.x.

Whitty, M. T., & Carville, S. T. (2008). Would I lie to you? Self-serving lies and other-oriented lies told across different media. *Computers in Human Behavior, 24,* 1021–1031. https://doi.org/10.1016/j.chb.2007.03.004.

Williams, S. S. (2001). Sexual lying among college students in close and casual relationships. *Journal of Applied Social Psychology, 31,* 2322–2338. https://doi.org/10.1111/j.1559-1816.2001.tb00178.x.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology, 14,* 1–59. https://doi.org/10.1016/s0065-2601(08)60369-x.

Deception Theories, Frameworks, and Approaches

CHAPTER 7

Implications of Some "Obvious Truths" for Building Theories of Deceptive Message Formulation and Production

John O. Greene, Kylie L. Geiman and Douglas E. Pruim

Researchers studying deceptive communication report that the phenomenon is actually quite common (e.g., Levine, Serota, Carey, & Messer, 2013; Serota, Levine, & Boster, 2010; Smith, Hancock, Reynolds, & Birnholtz, 2014), but despite its prevalence, there are not a great many efforts to explicate the nature of the cognitive mechanisms that underlie and give rise to deceptive behavior (but, e.g., see: Greene, O'Hair, Cody, & Yen, 1985; McCornack, 1997, McCornack, Morrison, Paik, Wisner, & Zhu, 2014; Walczyk, Harris, Duck, & Mulay, 2014). The central premise guiding this essay is that theorizing in this area can be advanced by taking into account a set of fundamental observations, things that "were right in front of us all along," but that have far-reaching implications for understanding how it is that people are able to convey to their interlocutors something other than what they consider to be true. The aims of this chapter, then, are, first, to

J. O. Greene (✉)
Brian Lamb School of Communication, Purdue University,
West Lafayette, IN, USA
e-mail: jgreene@purdue.edu

K. L. Geiman
SMG, Kansas City, MO, USA
e-mail: kgeiman@smg.com

D. E. Pruim
Krannert School of Management, Purdue University, West Lafayette, IN, USA
e-mail: dpruim@purdue.edu

sketch an agenda for theorizing about deceptive message production by high-lighting some "obvious truths" about the phenomenon whose implications for theory building have not been fully exploited. Following this, we examine the potential of one particular theory of message production, second-generation action assembly theory (AAT2; Greene, 1997, 2006), for explicating the cognitive processes underlying deceptive message behavior.

Some "Obvious Truths"

As indicated above, there are certain features of deceptive behavior that, once made explicit, have an "of course" character about them, but that can be seen to have far-reaching implications for guiding our thinking about how it is that deceptive messages come to be.

"Creation" Versus "Selection"

The first of our "obvious truths" is that there is more to the production of deceptive messages than selection from some repository of ideational content. It is convenient to think of message production as a process of *selection*—i.e., that there is some store of potential sayables (propositions, or what have you), and that message production involves selecting from this store the most appropriate (in some sense) content for expression. It is this sort of idea that one finds in Theo Hermann's (1983) conception of a "propositional base" of declarative knowledge from which some subset is selected as the "semantic input" for speech production.[1] The key point here is that message production, including deceptive message production, involves more than selection from some established repository of ideational content—people have a ready capacity for creating "ideas" that they have never heard, seen, or thought before (see Greene, 2006).[2] The challenge for theorists seeking to understand deceptive message production, then, is to explain how it is that each of us has this ability not simply to select, but to *create* new thoughts, utterances, and actions.

Individual Differences

A second key point is that people differ in their ability to deceive. To some extent, these individual differences in deception success may stem from relatively "static" features such as one's physiognomy (e.g., having an "honest face"). Other factors include a person's characteristic "communication style" or "demeanor"—that is, regardless of whether they are being honest or deceptive, some people simply come across as more sincere than others (see Levine, 2010; Levine et al., 2011). Beyond such factors, and most relevant to present concerns, is the potential impact of practice and skill on deception performance. We know, for example, that experienced salespersons are better able than the general populace to successfully perpetrate lies (see DePaulo, 1992). The upshot for theorists, then, is to develop descriptions of

the cognitive system that, in some way, accounts for the fact that behavior changes with practice.

Skill Repertoires

As an extension of the previous point, a third observation is that even within an individual, success at deception may vary as a function of domain and context. Thus, a person may be skilled at deception in some area, for example, feigning interest in the experiences of another, but relatively poor in some other aspect of social interaction. Similarly, a master poker player may have learned to suppress every "tell" at the casino, but be unable to successfully feign excitement over a disappointing gift from her children. Just as with the implications of practice for theories of deceptive message production, this latter point suggests that theories may need to accord a place for conceptions of context-specific "repertoires" of deception skills.

Verbal and Nonverbal Channels of Behavior Are Inextricably Intertwined

Our next observation is grounded in a reaction to the customary and convenient distinction between verbal and nonverbal channels of behavior. To be sure, various authors (see Andersen, 2008, Chapter 1) have gone to lengths to separate the two, and their arguments are informative. We believe, however, that it is also the case that the verbal and nonverbal features of message behavior are the products of an integrated action-production system.

From this root assumption, a fourth seemingly obvious point is that the behavioral resources at the deceiver's disposal are not restricted to the verbal realm. The idea here is not that there are various nonverbal cues to deception (see Bond & DePaulo, 2006; DePaulo et al., 2003). That's the way we typically think about the role of nonverbal cues—they accompany, and perhaps reveal, verbal acts of deception. Rather, the point is that the deceiver is able to employ nonverbal cues in perpetrating his or her deceptions, as when someone feigns a smile or, again, our crafty gambler maintains a "poker face."

The implication of this point is that theories of deception-enactment that are restricted to the verbal realm are necessarily inadequate. What we need are theories that accommodate both the verbal and nonverbal message channels and also explicate the nature of the mechanisms linking the two.

Nature and Role of Conscious Awareness

Yet another "obvious truth" is that there are certainly instances in which the deceiver is consciously aware of, and perhaps able to control, certain aspects of his or her verbal and/or nonverbal behavior. This, of course, is an idea commonly associated with Ekman and Friesen's (e.g., 1969) classic formulation, but what we have in mind regarding this point extends beyond that.

For the theorist, the deeper issue is accounting for (1) the nature of consciousness, and (2) the nature of the system by which conscious processes, or contents of consciousness, take your pick—it ultimately depends upon the specifics of one's theoretical formulation—drive overt verbal and nonverbal behaviors.

The obvious counterpoint to the fact that conscious monitoring and control of behavior transpires is that, at all times, some features of behavior proceed automatically, out of the realm of awareness and conscious control. But we would emphasize that the key point here is not to simply acknowledge the obvious; rather, the exigence for the theorist is to explain *how* this is so. And moreover, with a nod toward Kaplan's (1964) notion of an "esthetic norm" for theory building, to develop a model that addresses both the controlled and automatic properties of message behavior, deceptive and otherwise, in a seamless, integrated framework.

Spontaneity and Planning

Perhaps more so than the other "obvious truths" discussed here, this one may be so widely recognized that it need not be mentioned. Nevertheless, it is included here in the interest of providing a more comprehensive list of theoretical issues pertaining to deceptive message production. The point is simply that there are situations in which people produce deceptive messages "on the fly," with no prior preparation, and others where people give extensive thought to what they will say and do when perpetrating their deceptions—even to the point in some cases of overtly practicing those behaviors. And, although research on the effects of spontaneous- versus planned-lie manipulations indicates that those effects are somewhat limited, prior preparation is associated with shorter response latencies and reduced silent pausing (see DePaulo et al., 2003)—a pattern consistent with the more general research on the effects of planning on speech fluency (see Greene, 1995). For theorists, then, the implication is that functionalist models (see below) will need to incorporate some mechanism or system by which advanced planning is carried out, and, by extension, impacts subsequent behavior.[3]

Intrapersonal and Interpersonal Processes

Undoubtedly, one could go on extending this list of "obvious truths," seeming at each point to set ever more daunting conceptual challenges for the theorist, but we can bend this discussion in a different direction that introduces a whole new set of theoretical issues. To this point, the focus, and, by extension, the theorist's charge, has been upon intra-individual processes: e.g., not just selection but creation, nonverbal cues not just as an index of deception, but as a resource, and so on. These questions fall squarely within the cognitivist's wheelhouse. But, to concentrate on developing explanations for these phenomena alone opens the door to the familiar critique that cognitive

theories are too individualistic—that they give short shrift to the socially embedded nature of human thought and action.

At another level, then, there is the theoretical challenge of addressing the nature of deception as an *interpersonal* phenomenon. As Buller and Burgoon (1996, p. 204) noted, "Divorcing senders (deceivers) from receivers (detectors) draws attention to characteristics of individual actors rather than to properties of … interactions, to independent rather than to joint patterns of action…".

Here, again, there is something of a "this is news?" quality to our making a point of the interpersonal nature of deceptive episodes. The idea of deception as an interpersonal phenomenon is, of course, embedded in the very fabric of dramatic depictions of the parry-and-thrust of deceiver and potential deceit-detector since ancient times. But we have something different in mind here, and something different, too, from Goffman's (e.g., 1959) insight that self-presentation is typically a collaborative activity.

As we explicate more fully below, our point is to challenge theorists interested in the cognitive processes underlying deceptive message production to consider the socially interactive nature of thought and action. That is, rather than an exclusive focus on the information-processing activities of individuals, how might our understanding be advanced by efforts to explicate processes of mutual influence and processes of conjoint mentation?

And finally, as an extension of the preceding point, one last, set-the-theory-bar-higher, desideratum here is that theories addressing the intra-individual truths we began with and theories aimed at explicating mechanisms of interpersonal mutual influence ought not be separate, but rather should be elements of a seamless, integrated framework.

A Theory of Thought and Action: AAT2

Cognitive functionalist theories (see Greene & Dorrance Hall, 2013) seek to explain mental and overt behavioral phenomena by describing the system of (1) cognitive structures and (2) processes that give rise to and shape the behaviors of interest. In keeping with this perspective, AAT2 (see Greene, 1997, 2000, 2006) was developed as an effort to explicate verbal and non-verbal message behavior, again, by recourse to specification of the nature of the cognitive system underlying these actions. Despite the breadth and scope of the theory, its primary thrust can be conveyed by an examination of just two basic structural conceptions (i.e., procedural records and the output representation) and two processual notions (i.e., activation and assembly).

The Nature of Procedural Records

Central to the theory is the conception of a long-term memory (LTM) store of elemental units termed "procedural records." Two aspects of these structural entities are of particular note in the context of the current discussion.

First, a procedural record preserves relationships between features of (1) action, (2) outcomes associated with those action features, and (3) situational features under which that action–outcome relationship tends to hold. In effect, then, procedural records are "action – outcome – in situation" memory structures.

A second key point in explicating the nature of procedural records is that the action, outcome, and situational features in various records are represented in code systems that differ in levels of abstraction. Thus, some features are propositional in nature, others are lexically based (i.e., they correspond to the words of natural language), others are imagistic, some are motoric, and so on.

The upshot of this conception of elemental features which code just some small component of an individual's activity at any moment is a view of unfolding thought and action as a momentary collocation of a very large number of constituent features—some represented in quite abstract cognitive codes, others in mid-range representational systems, and still others that constitute the efferent (i.e., motor) programs that guide overt speech and movement.

The Activation Process

Given a conception of a memory repository containing a great many procedural records, there must be some process (or processes) by which the content of those memory representations is brought to bear in the production of behavior. Indeed, as noted above, according to AAT2, there are two such processes. The first, the "activation process" is essentially one of memory retrieval—that is, from the enormous store of procedural information in LTM, there must be some mechanism that serves to bring "appropriate" information to the fore.

AAT2 addresses this conceptual problem by positing that the action, outcome, and situational features of all procedural records are characterized at any moment by some level of activation which is heightened above resting levels when a person develops goals (or functional requirements, e.g., to express the concept "cat" one must identify an appropriate lexical item), or encounters either situational features or action features[4] that correspond, respectively, to the outcome, situational, and action features in that record. The result is that the LTM content that is most relevant to one's goals, requirements, and the situation at hand will tend to be that which is most highly activated above resting levels.

The Assembly Process

The AAT2 conception of activation is only one processual component of behavioral production: Although activation serves to retrieve relevant action features, it cannot, in and of itself, be a complete account of feature selection, or more generally, of output production. Consider that, as a result of

activation, an individual may have several highly activated candidate action features, only one of which will eventually find its way into overt behavior. As a simple example, at the level of word specifications, an individual may have at his or her disposal, "book," "tome," "volume," "paperback," "livre," and so on. The result of the activation process, then, is to retrieve a myriad of action features, at numerous levels of abstraction, and it is here that the assembly process comes into play. Assembly serves to integrate, or organize, activated action features into a (semi-) coherent (see Greene, 2000, 2006) representation of ongoing thought and action.

Just as one can think of activation as "retrieval," assembly in AAT2 is conceptualized as "coalition formation," and it is this process that constitutes the very heart of the theory. In AAT2, the time span of action-feature activation, and subsequent decay, is exceedingly rapid (i.e., in the vast majority of cases, just fractions of a second). However, the span of activation of any single feature is augmented when it "coalesces," or "fits with" other features, as for example, when the abstract propositional representation "this is delicious" coalesces with the lexical items, "this," "dish," "is" "eminently" "palatable," with an appropriate syntactic frame for ordering those words, and with motor-code specifications for both pronouncing those words and displaying facial cues of enjoyment. In contrast, those action features that do not find their way into coalitions (or that coalesce with few other features) are likely to quickly decay back to resting activation levels.

The Output Representation

The result of the assembly process is the second major structural component of AAT2, the "output representation." The output representation is, quite simply, the entire configuration of activated action features at any moment in time. The dynamic character of the output representation cannot be over-emphasized: Most of the features activated at any instant will quickly decay back to resting levels, "smaller" coalitions that are unable to recruit additional features will fragment and decay, abstract action specifications of what one is doing (or is "planning" to do) that do not coalesce with lower-level specifications for actually implementing those plans will not be manifested in behavior, and so on. At the risk of introducing too prosaic a metaphor, one can liken the dynamic nature of the output representation to watching the bubbles and turbulent swirls in a pot of boiling water (a primary shortcoming of the metaphor being that the processes unfolding in the output representation are considerably more complex than the intricate fluid dynamics of the roiling teapot on the stove).

Summary

To reiterate, the basic conceptual framework of AAT2 is defined by just two structural conceptions, procedural records and the output representation, and

two processes, activation and assembly. Obviously, there is more to be found in the finer details of the theory (see below), but even with this somewhat rudimentary sketch in place we are in a position to consider the implications of AAT2 for addressing the theoretical desiderata outlined in the first section of this chapter.

AAT2 AND PROCESSES OF DECEPTIVE MESSAGE FORMULATION AND PRODUCTION

"Creation" Versus "Selection"—Ideational Dynamics

At the outset, we noted that there is more to the production of deceptive messages than "selection"—that is, retrieving ideational content (and, presumably via some set of processes, expressing that content in overt behavior). Of course, it is possible that an individual may have at his or her disposal, stored, appropriate, misleading propositional content that only needs to be implemented (perhaps, e.g., as in the movies when one character instructs another on what to say when questioned by "the authorities"). Far more common, we suspect, are instances in which people exaggerate elements of their "factual base" of knowledge, omit elements of that knowledge store, bend certain details of what they think to be true, or, in some cases, simply create an account that has little relation to any actual event or experience.

Questions concerning how it is possible for people to go beyond simple selection of propositional content to be able to modify, or even create, ideational content are the focus of what Greene (2006) termed "ideational dynamics." As he notes (pp. 64–65), the essence of ideational dynamics is captured in the question: "*If thought drives talk, then what drives thought?*" (emphasis in the original). Recapitulation of the theorizing in that essay is not possible here, but in simplest terms, the "creative" character of thought and action is seen to be the inevitable consequence of the properties of the activation and assembly processes whereby established action features are combined (assembled) in always-novel configurations.

Individual Differences and the Effects of Practice

The second "obvious truth" introduced above was that people differ in their ability to engage in successful deceptions, and further, that expertise or practice almost certainly plays a role. The problem for theorists, then, is to specify the nature of the cognitive system that produces changes in behavior as a result of practice.

The relationship between practice and performance is obviously an issue that extends far beyond the relatively limited domain of deceptive message production. Indeed, the fact that practice leads to enhanced performance is a topic of long-standing focus for theorists, due in large measure to the fact that it is such a ubiquitous property of human behavior (see Adams,

1987; Lane, 1987; Newell & Rosenbloom, 1981; Proctor & Dutta, 1995). At a very general level of analysis, theoretical explanations for performance improvement as a result of practice typically invoke conceptions of (1) acquisition of new knowledge structures, (2) refining knowledge structures, and/or (3) strengthening memory structures (see Greene, 2003).

Each of these mechanisms is present in AAT2, but particular emphasis is given in the theory to the idea that combinations of elemental action features may, as a result of repetition, become stored together in memory. The ultimate result of such larger assemblages of action features is to reduce the likelihood that a person will encounter difficulties in coalition formation, either at the same level of symbolic representation or across representational systems, thereby resulting in the characteristic patterns of increased speed and performance quality associated with greater expertise (see Greene, 2011).

Skill Repertoires

As we noted at the outset, it is quite likely that some people will possess domain-specific deception abilities—that is, to be quite skilled at successful deception in certain contexts, but poor in others—thus, the need for theories to accord a place for context-specific "repertoires" of deception skills. From the perspective of AAT2, the context-specific nature of behavior, deceptive or otherwise, is simply the natural result of the "action – outcome – in situation" character of procedural records (along with the attendant notion that these records are strengthened with use). Thus, a "skill" that is evoked in one context (e.g., a car salesman lying to a buyer about a lemon) may not transfer to another situation (e.g., a car salesman lying to his wife about why he is coming home late) because the two are characterized by different sets of activating features.[5]

Verbal and Nonverbal Channels of Behavior Are Inextricably Intertwined

Our fourth point above centered on the need to extend theories of deceptive message production to address not only the verbal aspects of message behavior, and not simply to treat nonverbal cues as manifestations of verbally based cognitive machinations (e.g., indicants of cognitive load), but rather to treat nonverbal features of behavior as deception resources in their own right, and, beyond this, to develop theories that meld verbal and nonverbal message productions in an integrated fashion.

From the perspective of AAT2, verbal and nonverbal features of action are products of the same system of structures and processes—that is, although the symbolic representation of action features of certain aspects of verbal behavior will differ from those coding action features linked to nonverbal behavior (see Greene & Graves, 2007), the same processes of activation and assembly drive both aspects of behavior as well as the interaction between the two.[6]

Nature and Role of Conscious Awareness

In the first section of this chapter, we noted that there are situations when (1) people consciously monitor and control certain aspects of their behavior, that (2) even when some aspects of ongoing behavior are under conscious control, other aspects of their behavior are carried out automatically, out of conscious awareness, and (3) the behavioral specifications of which one is aware tend to play a major role in driving overt behavior. Again, the task of the theorist is to go beyond such "surface" observations to address the "how" and "why" of these phenomena.

From the perspective of AAT2, the activation and assembly processes (and hence, behavioral production) churn away, automatically and quite apart from conscious awareness. But, the theory does accord a key functional role in behavioral production to conscious awareness. The theory holds that coalitions that are characterized by high levels of activation, and that persist in that state for some period of time, are those that enter conscious awareness. As an extension of this point, coalitions whose contents are represented in more abstract code systems (and whose activating conditions persist for longer periods of time), and those that have larger numbers of constituent action features, are more likely to enter conscious awareness (Greene, 2006). The effect of consciousness, in turn, is to (1) bring "executive processes" (e.g., editing, planning, rehearsal) into play, (2) thereby augmenting the activation of relevant coalitions, (3) increasing the likelihood that those coalitions will martial lower-level action specifications, and (4) be manifested in overt behavior (see Greene, 2006). Again, however, the theory posits that the vast bulk of what goes on via activation and assembly proceeds quite automatically, out of conscious awareness: Coalitions are formed, recruit efferent specifications (thereby shaping overt behavior), and decay back to resting levels.

Among the implications of the conception of coalition formation in AAT2 is that behavioral production tends to be top-down (i.e., that abstract action specifications drive lower-level behavior), but the theory also makes explicit that this need not necessarily be so—that there are conditions under which lower level action specifications (e.g., a hand gesture) can drive higher level mentation. Moreover, although automatic and consciously controlled behaviors are treated as the product of a single, integrated system, the theory also addresses the fact that there are conditions that will give rise to discontinuities between various elements of the output representation (see Greene, 2006).

Spontaneity and Planning

In the introductory section of this chapter, we noted that there are occasions when people plan in advance their deceptive message behavior, and that such advance planning tends to increase message production fluency. The task for

theorists, then, is to explicate the nature of planning processes and the mechanisms by which prior preparation comes to impact subsequent behavior. In simplest terms, AAT2 treats planning as assembly of portions of the output representation in advance, and one effect of such prior assembly is to obviate the need for "on-line" assembly of those elements, thereby speeding the process of overt behavioral production.

What may be more noteworthy is that AAT2 incorporates no special "planning" or "rehearsal" mechanism. As Greene (2000, p. 152) states:

> [P]lans are viewed as coalitions – typically coalitions that include relatively abstract imaginal, propositional, and verbal constituents. Like all coalitions, these complexes are subject to decay (resulting in their passing out of consciousness, and perhaps complete loss), incorporation of new features and subcoalitions (resulting in their transformation "on the fly"), and reinstatement in consciousness if they once again come to exceed some activation threshold. Thus, the fluid and malleable character of plans is seen to arise as a natural consequence of the dynamics of activation and assembly.

Intrapersonal and Interpersonal Processes

Our final point from the first section of this chapter was to encourage theorists, and particularly those theorists who take a cognitive, functionalist approach to thinking about message production processes, to consider interpersonal, as well as intrapersonal, mechanisms at work in deceptive episodes. Beyond this, we suggest that it would be desirable to incorporate intraindividual and inter-individual processes in a seamless, integrated fashion.

Our own thinking on this point is shaped by the theory of transcendent interactions (TTI; Greene & Herbers, 2011), which is, in fact, an extension of AAT2 to the realm of interpersonal mutual influence processes. The project of TTI is to explicate processes of "conjoint mentation"—to address the socially interactive nature of thought and action (or as Greene [2006, p. 73] noted, to understand how it is that "interaction drives thought").

As the name of the theory suggests, the focus of TTI is upon, seemingly rare, interpersonal experiences of complete immersion, insight, and connection. But such experiences are held to define the ultimate endpoint of a continuum of along which all interactions can be arrayed, according to how near or far they fall from the "transcendent ideal." Thus, we would posit that the TTI framework has implications for understanding deceptive episodes—both those that unfold "without a hitch" and those that collapse completely. As an example, among the propositions of TTI is the idea that one's interlocutor may create assembly difficulties (as, e.g., by questioning the authenticity of an account), but also that that same conversational partner can present solutions to assembly difficulties (e.g., suggesting a plausible explanation for the gaps in one's story).

Conclusion

Survey of the Table of Contents of this volume reveals an impressively large and diverse array of scholarly treatments related to the general topic of "deception." Indeed, the breadth of the issues reflected here leads one to suspect that it is unlikely that any particular individual, even one completely focused on deception as his or her exclusive area of scholarly interest, could possibly keep tabs on all the work that is being done in the area. Rather, it seems more likely that the sub-discipline of deception research may have become marked by the proverbial "silos" of relatively isolated programs of study. To the extent that this is true, volumes such as this provide an opportunity to survey the sweep of current thought in the area. In that spirit, this chapter is intended to pose a set of conceptual challenges pertaining to cognitive functionalist approaches to theorizing about deceptive message formulation and production, and to provide an overview, a "primer" of sorts, of one particular framework for addressing those issues.

AAT2 falls squarely within the cognitive functionalist approach to theorizing about message production processes (including deceptive behavior), but it does differ in some important respects from other functionalist treatments (e.g., Gombos, 2006; McCornack, Morrison, Paik, Wisner, & Zhu, 2014; Walczyk, Harris, Duck, & Mulay, 2014; see also Greene, 2014), including in its characterization of "executive processes," such as planning and rehearsal, and the fact that the theory does not invoke a conception of "processing capacity" (see Greene, 1997). Grappling with these issues, then, can be seen to represent yet one final challenge for future theorizing.

Notes

1. To be fair, Hermann (1983) does acknowledge that memory content can be "enriched and elaborated" (p. 36), but the nature of these processes are not explicated (see Greene & Graves, 2007). For a noteworthy application and extension of Hermann's theorizing to deceptive message production where processes of manipulation of propositional content are explicitly addressed, see McCornack, Morrison, Paik, Wisner, and Zhu (2014).
2. It is also the case that, while every person routinely and as a matter of course, creates novel ideational content, people do differ in the facility with which they are able to do so (see Greene, Morgan, McCullough, Gill, & Graves, 2010; Morgan, Greene, Gill, & McCullough, 2009).
3. With regard to the nature of planning, there is, of course, a long tradition of work on this topic, both within the field of communication and in cognate disciplines (see Berger, 1997; Friedman, Scholnick, & Cocking, 1987).
4. In point of fact, the idea that observing actions on the part of an interlocutor could serve to activate one's own action representations was not introduced until the development of Greene and Herbers' (2011) theory of transcendent interactions.
5. See Greene and Geddes (1993) for a detailed exposition of skill transfer in action assembly theory.

6. To assert that verbal and nonverbal features of behavior are the product of a single, integrated message production system should not be taken to imply that the association between the two is seamless. Greene (2000) makes a central point of addressing how difficulties in coalition formation may lead to a lack of continuity and consistency within and across hierarchical levels of behavioral specification.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin, 101,* 41–74. https://doi.org/10.1037/0033-2909.101.1.41.

Anderson, P. A. (2008). *Nonverbal communication: Forms and functions* (2nd ed.). Long Grove, IL: Waveland.

Berger, C. R. (1997). *Planning strategic interaction: Attaining goals through communicative action.* Mahwah, NJ: Lawrence Erlbaum.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Human Communication Research, 6,* 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111,* 203–243. https://doi.org/10.1037/0033-2909.111.2.203.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32,* 88–106. https://doi.org/10.1080/00332747.1969.11023575.

Friedman, S. L., Scholnick, E. K., & Cocking, R. R. (Eds.). (1987). *Blueprints for thinking: The role of planning in cognitive development.* New York, NY: Cambridge University Press.

Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday Anchor.

Gombos, V. A. (2006). The cognition of deception: The role of executive processes in producing Lies. *Genetic, Social, and General Psychology Monographs, 132,* 197–214. https://doi.org/10.3200/MONO.132.3.197-214.

Greene, J. O. (1995). Production of message in pursuit of multiple social goals: Action assembly theory contributions to the study of cognitive encoding processes. In B. R. Burleson (Ed.), *Communication yearbook 18* (pp. 26–53). Thousand Oaks, CA: Sage.

Greene, J. O. (1997). A second generation action assembly theory. In J. O. Greene (Ed.), *Message production: Advances in communication theory* (pp. 151–170). Mahwah, NJ: Lawrence Erlbaum.

Greene, J. O. (2000). Evanescent mentation: An ameliorative conceptual foundation for research and theory on message production. *Communication Theory, 10,* 139–155. https://doi.org/10.1111/j.1468-2885.2000.tb00184.x.

Greene, J. O. (2003). Models of adult communication skill acquisition: Practice and the course of performance improvement. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 51–91). Mahwah, NJ: Lawrence Erlbaum.

Greene, J. O. (2006). Have I got something to tell you: Ideational dynamics and message production. *Journal of Language and Social Psychology, 25,* 64–75. https://doi.org/10.1177/0261927X05284481.

Greene, J. O. (2011). Communication skill acquisition. In W. Donsbach (Ed.), *International Encyclopedia of Communication*. Malden, MA: Blackwell. https://doi.org/10.1002/9781405186407.

Greene, J. O. (2014). Just looking around. *Journal of Language and Social Psychology, 33,* 398–404. https://doi.org/10.1177/0261927x14536051.

Greene, J. O., & Dorrance Hall, E. (2013). Cognitive theories of communication. In P. J. Schulz & P. Cobley (Eds.), *Handbooks of communication science, Vol. 1: Theories and models of communication* (pp. 181–198). Berlin: Mouton de Gruyter.

Greene, J. O., & Geddes, D. (1993). An action assembly perspective on social skill. *Communication Theory, 3,* 26–49. https://doi.org/10.1111/j.1468-2885.1993.tb00054.x.

Greene, J. O., & Graves, A. (2007). Cognitive models of message production. In D. R. Roskos-Ewoldsen & J. L. Monahan (Eds.), *Communication and social cognition: Theories and methods* (pp. 17–45). Mahwah, NJ: Lawrence Erlbaum.

Greene, J. O., & Herbers, L. E. (2011). Conditions of interpersonal transcendence. *International Journal of Listening, 25,* 66–84. https://doi.org/10.1080/10904018.2011.536472.

Greene, J. O., Morgan, M., McCullough, J., Gill, E., & Graves, A. R. (2010). A phrase well turned: Creative facility in narrative production. *Communication Studies, 61,* 118–134. https://doi.org/10.1080/10510970903400329.

Greene, J. O., O'Hair, H. D., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. *Human Communication Research, 11,* 335–364. https://doi.org/10.1111/j.1468-2958.1985.tb00051.x.

Hermann, T. (1983). *Speech and situation: A psychological conception of situated speaking*. Berlin: Springer.

Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. New York, NY: Harper & Row.

Lane, N. E. (1987). *Skill acquisition rates and parameters: Issues and training implications*. New York, NY: Spring.

Levine, T. R. (2010). A few transparent liars. In C. T. Salmon (Ed.), *Communication Yearbook 34* (pp. 40–61). New York, NY: Routledge.

Levine, T. R., Serota, K. B., Carey, F., & Messer, D. (2013). Teenagers lie a lot: A further investigation into the prevalence of lying. *Communication Research Reports, 30,* 211–220. https://doi.org/10.1080/08824096.2013.806254.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., et al. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research, 37,* 377–403. https://doi.org/10.1111/j.1468-2958.2011.01407.x.

McCornack, S. A. (1997). The generation of deceptive messages: Laying the groundwork for a viable theory of interpersonal deception. In J. O. Greene (Ed.), *Message production: Advances in communication theory* (pp. 91–126). Mahwah, NJ: Lawrence Erlbaum.

McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology, 33,* 348–377. https://doi.org/10.1177/0261927X14534656.

Morgan, M., Greene, J. O., Gill, E. A., & McCullough, J. D. (2009). The creative character of talk: Individual differences in narrative production ability. *Communication Studies, 60,* 180–196. https://doi.org/10.1080/10510970902834908.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Lawrence Erlbaum.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance.* Thousand Oaks, CA: Sage.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36,* 2–25. https://doi.org/10.1111/j.1468-2958.2009.01366.x.

Smith, M. E., Hancock, J. T., Reynolds, L., & Birnholtz, J. (2014). Everyday deception or a few prolific liars? The prevalence of lies in text messaging. *Computers in Human Behavior, 41,* 220–227. https://doi.org/10.1016/j.chb.2014.05.032.

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology, 34,* 22–36. https://doi.org/10.1016/j.newideapsych.2014.03.001.

Got Theory? Multitasking, Cognitive Load, and Deception

Tripp Driskell and James E. Driskell

The study of deception is relevant to a broad variety of fields, including psychology, communication, political science, law and forensics, sociology, and many other domains. Despite the widespread interest in this topic across various disciplines, it is a somewhat unique topic in that the goal of much research in this area is to make human performance *worse*. That is, one implicit goal of research on deception is to improve the detection of deception or to examine conditions under which liars fail. In other words, deception researchers endeavor to create situations in which people are unsuccessful at lying. In addition to impeding the performance of liars, researchers also want to maintain, within reason, the performance of truthtellers. This is done to maximize differences between truthtellers and liars in the hope that cues to deceit will be more salient and thus easier to distinguish. Moreover, the deception literature makes efforts to understand factors that are related to deception ability, not to promote high performance, but to better understand lies so that they can be more easily caught.

Despite its uniqueness, the deception literature has many similarities with many other fields of study. The most prominent of which is the existence of competing theoretical perspectives. There is a fault line that runs through the deception community that divides the field into two broad camps: an emotion camp and a cognitive camp. The emotion camp argues that liars may experience emotions (e.g., guilt, anxiety, duping delight) when telling a lie, and these emotions, in turn, are diagnostic of deception. However, it

T. Driskell · J. E. Driskell (✉)
Florida Maxima Corporation, Orlando, FL, USA
e-mail: JDriskell@Rollins.edu

is difficult to determine under what conditions liars experience which, if any emotions. For instance, one of the earliest researchers to study deception, Hugo Munsterberg (1908), noted that more experienced liars may not be afflicted by intense emotional responses to lying. Moreover, emotional reactions are not only restricted to liars. In many cases telling the truth may evoke the same types of emotions (e.g., concern about not being believed). In short, criticisms of the emotion-based approach have stemmed from its inability to reliably predict differences between lying and telling the truth (Vrij, Fisher, Mann, & Leal, 2008).

The cognitive camp argues that cues to deception can best be predicted by a cognitive load approach. Generally speaking, this approach posits that it is more cognitively demanding to fabricate a lie than to tell the truth and, in turn, the resultant increase in cognitive demand may be reflected in observable behaviors. One of the early formulations of the cognitive load approach, advanced by Zuckerman, DePaulo, and Rosenthal (1981), states that lying is more cognitively complex because (a) fabricating details is more difficult than telling the truth, (b) liars must work to avoid logical inconsistencies, and (c) their stories cannot contradict what the listener knows. A more recent and detailed explanation of the cognitive load approach offered by Vrij and colleagues (e.g., Vrij, Fisher, & Blank, 2017; Vrij et al., 2008; Vrij, Granhag, Mann, & Leal, 2011; Vrij, Mann, et al., 2008) suggests that lying is more cognitively demanding than telling the truth because:

1. Fabricating a lie is cognitively demanding.
2. Liars are less likely to take their credibility for granted and are thus more likely to monitor and control their own behavior to appear honest.
3. Liars are less likely to take their credibility for granted and are thus more likely to monitor the listener to determine if they are succeeding at their lie.
4. Liars may be preoccupied with the task of reminding themselves to act and role-play.
5. Lying requires a justification (e.g., lying to avoid arrest), which is cognitively demanding.
6. Liars need to suppress the truth while they are lying.
7. Lying is intentional and deliberate, which is cognitively demanding.

Moreover, Vrij (2015; Vrij et al., 2017) argues that the cognitive load approach offers a further advantage in that it suggests specific interventions that may magnify the difference in cognitive load between liars and truthtellers and thus lead to greater discrimination between the two. Interventions that have been proposed include (a) imposing cognitive load to make the task more difficult, (b) asking anticipated questions that may result in greater

cognitive demand on the liar relative to the truthteller, and (c) encouraging greater information to be provided, given that a fabricated story generally lacks detail and providing this requires greater effort.

Although the cognitive load approach to lie detection has considerable empirical support (Vrij, 2015), the precise theoretical rationale, or the degree to which this approach is linked to existing theories of cognition, is not elaborated thoroughly. This is not particularly surprising given that (a) the deception literature (DePaulo et al., 2003; Vrij et al., 2011) is distinct from the multitasking and cognitive load literatures (Lavie, 2010; Salvucci & Taatgen, 2008; Wickens, 2002; and others) from which they stem, (b) primary-level research on deception and on cognitive load is conducted in different subfields of psychology and related disciplines, and (c) this research is generally published in distinct and separate journals. Therefore, the primary objective of this chapter is to better tie the cognitive load approach to deception to the broader cognitive load literature on which it is based. To accomplish this objective, we will review theoretical approaches to multitasking and cognitive load while relating these approaches to deception where applicable. This review will also allow us to provide specific propositions for examining lie detection that emerge from these approaches. These propositions are provided to stimulate future research efforts.

Theories of Multitasking and Cognitive Load

It should be apparent from the seven-point description of the cognitive load approach presented above that lying is not a simple task. More importantly, it should be apparent that the process of lying is multifaceted and may become more or less complex given the degree to which each of the seven principles occur simultaneously during a lie. For instance, lying may be more difficult when a liar monitors both his/her *and* a listener's behaviors, than if he/she was merely monitoring their own behaviors. To execute a successful lie, liars need to be able to perform multiple tasks well. It is when the cognitive faculties required to execute multiple tasks are lacking that liars get into trouble. This is the basic premise underlying the concept of multitasking.

Multitasking can be defined as managing and executing more than one task at a time (Salvucci & Taatgen, 2008) or performing multiple tasks within a limited time period (Fischer, Morrin, & Joslyn, 2003). This construct is related to several terms including dual-task performance, task load, and workload.[1]

Our ability to perform multiple tasks simultaneously is a notable feat of human cognition; however, this ability is imperfect. Researchers, more specifically resource or capacity theorists, have spent decades explaining the performance decrements associated with carrying out multiple tasks. This research has largely been conducted using a psychological refractory period

(or overlapping tasks) paradigm (see Smith, 1967). The results from these studies have been used as evidence of two primary classes of theories: bottleneck (task-switching) theories and resource theories (cf. Meyer & Kieras, 1997; Pashler, 1994). These classes of theories are briefly reviewed below.

Bottleneck (Task-Switching) Theories

Bottleneck, or task-switching, theories are said to originate from the work of Telford (1931) and propose "that there are certain stages of processing (constituting a bottleneck) that cannot be performed simultaneously on more than one input" (Pashler, 1984, p. 358). In the dual-task paradigm, an individual is required to perform two concurrent tasks. If both tasks require similar processing structures, then the tasks are queued, resulting in one task being carried out and then the other. As a result, when individuals attempt to perform multiple tasks simultaneously there may be a delay in the primary and/or secondary tasks. This class of theories has witnessed an evolution from single-channel hypotheses (e.g., Craik, 1948; Telford, 1931; Welford, 1967) to early- (Broadbent, 1958) and late-selection (Deutsch & Deutsch, 1963; Keele, 1973; Pashler, 1984) structural bottleneck models (Meyer & Kieras, 1997). Although these theories share the same assumption that tasks must be performed in sequence, they differ in regard to where they propose that the bottleneck (or bottlenecks) occurs (e.g., in perception or in response). Although bottleneck theories still garner support, they have largely given way to resource theories.

Resource Theories

A central assumption of resource theories is that individuals have a limited amount of cognitive resources that can be flexibly allocated among tasks (Baddeley, 1986; Baddeley & Logie, 1999). However, successful task allocation depends on the types of resources required to perform a task. If an individual is executing two tasks that require similar cognitive resources, then the resources need to be divided between the tasks. Therefore, fewer cognitive resources are available to perform each task than there would be if a single task was performed. Moreover, task performance depends on task prioritization. That is, if a task is prioritized, then limited attentional resources will be allocated to the primary task and secondary task performance suffers (Kahneman, 1975; Mcleod, 1977; Neisser & Becklen, 1975). Additionally, if both tasks have equal priority, errors may be observed on both tasks.

The primary difference between bottleneck theories and resource theories is that tasks can be performed simultaneously in resource theories. However, it should be noted that both these classes of theories are not mutually exclusive. That is, as Navon and Miller (2002) point out, "resource theory actually subsumes the single-bottleneck notion, where a serial bottleneck is regarded

as a special type of resource" (p. 195). Resource theories range from unitary resource theory (e.g., Kahneman, 1973; Moray, 1967) to multiple resource theory (Meyer & Kieras, 1997; Navon & Gopher, 1979; Salvucci & Taatgen, 2008; Wickens, 1980).

In brief, research within the dual-task paradigm indicates that concurrent tasks interfere with one another because of the increased demands on limited attentional and processing capacity. This would suggest that when liars[2] carry out multiple concurrent tasks during a lie, errors may occur during task performance that could reveal their lie (e.g., response latency). However, this will not always be the case. If the demands of multitasking are negligible—that is, cognitive resources are not fully taxed—then performance may not suffer. Consequently, it is important to determine when task demands are taxing enough to impair a liar's performance. To determine this, we need to go beyond the high level of abstraction presented above and delve deeper into a theory of dual-task performance. This will afford us the ability to extract more discrete propositions. Below we present Wickens' Multiple Resource Theory (MRT: Wickens, 1984, 1991, 2005; Wickens & Liu, 1988). This theory was chosen for several reasons. First, MRT is a type of resource theory which, as Pashler (1994) notes, is "probably the most accepted way to think about dual-task interference" (p. 221). Second, MRT is very detailed and thus allows us to draw more specific propositions. Third, it has largely passed the litmus test of any psychological theory: longevity, empirical scrutiny, and predictive power (see Wickens, 2002). And fourth, Wickens (2008) notes that MRT should be neurophysiologically plausible (i.e., MRT's dimensions have parallels in brain anatomy). This is relevant due to the increased interest in the neural substrates of deception (see Abe, 2011).

Multiple Resource Theory

Multiple resource theory aims to understand the degree to which multitasking affects time-sharing ability and ultimately, performance (Wickens, 2008). Wickens' MRT affords detailed predictions about how well each task in a multitask set will be performed in combination, relative to how each is performed alone. Moreover, MRT provides a detailed understanding of the mechanisms by which decrements in performance are produced. According to MRT, three primary factors determine multitasking performance. They are task demand, resource overlap, and resource allocation. These are described in detail below.

Task demand. Demand refers to the difficulty of a task set. We can conceptualize task demand as lying on a continuum ranging from simple to difficult. This task demand continuum determines, in part, multitasking ability. In short, two simple tasks can be performed simultaneously without a performance decrement because they do not exhaust attentional resources. Wickens (2002) refers to these types of tasks as "residual capacity" or "data limited."

On the other hand, as task demand increases, concurrently performed tasks compete for limited cognitive resources, and at a certain point cognitive resources are exhausted and performance begins to suffer. Tasks are said to be "resource limited" when they demand all available resources (Wickens, 2002). Consequently, multitasking performance is a function of task difficulty.

Task demand[3] is most closely associated with the concept of workload or cognitive load. It is also most closely associated with the intervention proposed by Vrij, Fisher, et al. (2008) and Vrij et al. (2011) to *impose cognitive load* in interview settings, which aims to increase the demands of lying in order to evoke more salient cues to deception. Increasing the demands of a task can be accomplished by heightening intrinsic or extrinsic task difficulty. Intrinsic task difficulty is related to a variety of factors including: (a) task structure, including the rate at which information is presented and the length of time the task must be performed, (b) the requirements for speed or accuracy (i.e., almost any task can be made more difficult by increasing the demands for additional speed or accuracy), and (c) task complexity such as the number of alternative choices to be considered and the ambiguity and novelty of the information presented (Huey & Wickens, 1993). Extrinsic factors represent characteristics unrelated to the task itself, including stress, motivation, and other external demands. This exposition suggests the following propositions.

Proposition 1: Lying can be made more difficult by manipulating the intrinsic difficulty of a task. In fact, this is what deception researchers have recently been doing. However, most studies have focused on increasing task complexity. For example, researchers have had participants maintain eye contact with interviewers (Vrij, Mann, Leal, & Fisher, 2010) or to recall events in reverse order (Vrij, Mann, et al., 2008). Although these studies have proven valuable, there are additional ways researchers can increase task complexity (see following section on Task Moderators).

Proposition 2: Lying can be made more difficult by manipulating extrinsic task demands. Deception researchers regularly call for more ecologically valid paradigms in which to conduct research. In laboratory-based studies, it is difficult to simulate the stressors present in real-world lies (e.g., the consequences of failure). One reason that approximating the stressors of real-world lies is desirable is because theoretically, liars should perform worse under these conditions and are thus easier to detect. Deception research has substantiated this claim (e.g., ten Brinke, Porter, & Baker, 2012; Vrij & Mann, 2001). There is also clear evidence from the stress literature that performance suffers under stressful conditions (see Driskell & Salas, 1996; Hancock & Szalma, 2012). Although it has proven difficult to create an ecological valid experimental paradigm, steps can be taken to manipulate extrinsic demand factors in both laboratory and applied settings (see Driskell, Mullen, Johnson, Hughes, & Bachelor, 1992).

Caveat: Well-learned tasks mitigate task difficulty. Specifically, practice can lead to task automaticity. Automatic tasks require less resources to be executed (Wickens, 2002; See Proposition 11).

Resource overlap. A fundamental tenant of MRT is that individuals possess multiple "pools" of resources and therefore, multitasking performance is dependent on which "pools" a task demands. According to MRT, a task can be characterized by demands on either dichotomous level of four resource dimensions: (a) processing stages (perception and working memory vs. response), (b) processing codes (verbal vs. spatial), (c) processing modalities (auditory vs. visual), and (d) visual channels (focal vs. ambient). If we imagine a 4-dimensional cube representing these four resource dimensions, we note that any task may occupy one or more of the cells in the cube. If two tasks demand different cells in the cube, then interference isn't likely to occur and multitasking performance is enhanced. However, if two tasks occupy the same cell, dual-task interference is likely to occur because they vie for the same limited resources.

Multiple resource theory identifies three *processing stages*: perception, cognition (e.g., working memory), and response (e.g., action). According to Wickens et al. (2016), perceptual and cognitive activities demand similar attentional resources. Moreover, the resources that these activities require appear to be different from the resources that response execution requires. For example, Shallice, McLeod, and Lewis (1985) found evidence that the resources underlying speech recognition and speech production (a perceptual and response activity respectively) were distinct.

Proposition 3: Two tasks both demanding either (a) perceptual or cognitive processes (e.g., working memory, decision making, information integration) or (b) response processes will interfere with each other to a greater extent than will two or more tasks, one of which requires perceptual or cognitive processes and the other requiring response processes.

The second dimension of MRT, *processing codes,* refers to whether information is processed spatially or verbally. Conceptually like processing stages, spatial activities demand different resources than verbal activities (Wickens, 2008). In combination with processing stages, MRT indicates that spatial and verbal activities use different processes, whether operating in perception, cognition, or response (Wickens, 1980).

Proposition 4: If multiple tasks concurrently tax verbal cognition or concurrently tax processing, multitasking is more difficult. Conversely, dual-task interference is less if one task is responded to verbally and the other task requires a manual response.

The third dichotomous dimension of MRT, *perceptual modalities*, represents a distinction between processing information via the visual or the auditory channels. As Wickens (2002) notes "cross-modal time-sharing is better than intra-modal time-sharing" (p. 164). In other words, it is easier to divide one's attention between a visual task and an auditory task than to divide one's

attention between two tasks that demand visual or auditory processing. To provide an example, monitoring an interviewer's behavior and listening to an interviewer's questions requires both visual and auditory channels. However, intra-modal time-sharing would occur if, for instance, questions were presented visually or the interviewee was instructed to pay attention to a certain visual stimulus (e.g., maintain eye contact). That is, requiring an interviewee to pay attention to a certain visual stimulus as well as (simultaneously) monitoring the interviewer's behavior would lead to greater cognitive load.

Proposition 5: Tasks that require liars to engage in intra-modal visual or auditory time-sharing may complicate their ability to deceive.

The fourth dimension, *visual channels*, was a latter add-on to MRT. This dimension represents a dichotomy in visual processing: focal versus ambient vision. Focal vision aids in distinguishing fine details and pattern recognition, while ambient vision aids in peripheral activities (e.g., sensing orientation; Wickens, 2002). Again, multiple tasks can be completed with minimal interference if one task demands focal vision (e.g., catching a football) and the other demands ambient vision (e.g., running a route). To provide a more relevant example, maintaining eye contact and monitoring interviewer behaviors require focal vision. Thus, if an interviewee attempted to perform both simultaneously, a dual-task performance decrement on one or both tasks would be likely to occur.

Proposition 6: Tasks that require the same visual channel concurrently (i.e., both focal or both ambient vision) will be more difficult to perform than a task that requires the use of separate visual channels (i.e., focal vision only, ambient vision only, or focal and ambient vision).

Resource allocation. The final factor, *resource allocation*, describes the allocation of resources between tasks. Resource allocation is performed by the executive control system (Baddeley, 1983, 1995). The executive control system is charged with task prioritization, attention switching, and interruption management, which in and of themselves demand cognitive resources (Wickens et al., 2016). In other words, resource allocation competes with some of the same resources that task demand and resource overlap compete for.

The manner in which limited attentional resources are allocated has a significant impact on how multiple tasks are performed and which tasks are most likely to suffer. In experimental research, primary and secondary tasks are usually prescribed. However, in real-life situations, tasks may not be overtly prescribed and may vary in importance across the life span of a task. For example, when lying, it is suggested that individuals are more inclined to monitor an interviewer to determine if they are getting away with their lie. This can create dual-task interference (e.g., by both monitoring the interviewer's behaviors while listening and responding to questions). When monitoring becomes the primary task, we are more likely to see errors on the other secondary tasks (when cognitive demand exceeds limited resources).

However, a liar may only monitor interviewer behavior when they feel they have the resources to do so. Moreover, if the liar feels secure that the lie is succeeding, they may quit monitoring altogether and allocate resources to other deception relevant tasks.

Proposition 7: Errors during deception are more likely to occur on secondary task performance. However, which specific task is secondary depends on how resources are allocated. It is also possible that two tasks will both share priority, in which case both tasks will suffer equally. As a result, when and where errors occur may vary across time. This suggests that there will be a "critical response period" in which errors are more likely or less likely to occur, dependent on how attentional resources are allocated. Related to Signal Detection Theory (Green & Swets, 1966), this can help in identifying when and where to look for errors related to deception (i.e., to better distinguish the signal from the noise).

Moderators and Qualifications

The following section examines moderators of multitasking ability. These moderating variables will have a direct impact on deception ability and the cues that are exhibited during deception.

Intelligence

Multitasking performance is based largely on the amount of cognitive resources one can devote to task execution. It stands to reason that individuals with greater stores of resources will be better at multitasking than individuals with fewer stores of resources. As research indicates, there seems to be a strong relationship between multitasking and measures of working memory capacity and general intelligence (Brookings, 1990). This notion is also addressed within the deception literature (Vrij, 2008). For instance, Vrij, Granhag, and Mann (2004) note that less intelligent people find it more difficult to lie. However, recent research failed to demonstrate a correlation between deception ability and intelligence (Wright, Berry, & Bird, 2012). It is possible that the cognitive demands in this study were not sufficient enough to produce an effect for intelligence. In other words, task performance may not have required the full allocation of cognitive resources.

Proposition 8: Less intelligent populations will have more difficulty fabricating and executing a lie. As a result, cues to deception (especially cognitive load cues) will be more pronounced in these populations.

Caveat: The lie that is told must be taxing enough to consume the cognitive resources of the less intelligent population. If this is achieved, there will be a greater performance deficit in the less intelligent population than in the more intelligent population.

Age

It is claimed that with age comes wisdom, which is a comforting thought to many of us. Unfortunately, research indicates that age is also accompanied by a variety of cognitive issues. For example, research has demonstrated that older people may have less capacity for coordinative processing of multiple sources of information (Mayr, Kliegl, & Krampe, 1996) and that attentional flexibility and the ability to avoid distractions decline with age (Fisk & Rogers, 2007; Gazzaley, Cooney, Rissman, & D'Esposito, 2005). Meta-analyses examining the relationship between age and dual-task performance have found a clear impairment associated with age (Riby, Perfect, & Stollery, 2004; Verhaeghen, Steitz, Sliwinski, & Cerella, 2003).

Proposition 9: Deception ability will decline with age. The available research suggests that older individuals will have greater difficulty with multitasking that requires a large degree of controlled processing (Riby et al., 2004).

Caveat: Cognitive deficits associated with aging are task dependent. The meta-analysis conducted by Riby et al. (2004) demonstrated dual-task impairment on controlled processing tasks and no such impairment on automatic processing tasks. Automatic processing tasks require less attention and consume less cognitive resources. As a result, we would not expect to see an age effect on deception ability for simple or well-learned tasks.

Neuroticism and Social Anxiety

Third, research suggests that certain personality characteristics may influence multitasking ability. Specifically, research has demonstrated that people who score high on neuroticism have more difficulty multitasking than those who score low on neuroticism. For example, Poposki, Oswald, and Chen (2009) proposed that people high in neuroticism (reflecting negative emotionality and anxiety) perform more poorly under highly stimulated conditions such as multitasking. Their findings demonstrated a significant negative relationship between neuroticism and multitasking performance. Related to neuroticism, social anxiety has been linked to deception ability (Vrij, Akehurst, Soukara, & Bull, 2002, 2004). Specifically, a negative relationship between neuroticism and Criteria-Based Content Analysis (CBCA: a content analytic approach to detecting deceptive communications) score was found. Interestingly, this effect was only found for liars and older adults (Vrij et al., 2002).

Proposition 10: Individuals that score high on neuroticism—which reflects negative emotionality and anxiety—will have poorer deception ability than those who score low on neuroticism. This proposition is linked to the research presented above, as well as self-presentation theory (see Baumeister, 1982; DePaulo et al., 2003). In short, liars who are high on neuroticism and/or are socially anxious will be more likely to be concerned with presenting a specific impression (DePaulo & Tang, 1994), and consequently will devote more cognitive resources to managing that impression and the emotions evoked by

the consequences of not being able to do so. This, in turn, may cognitively burden the individual resulting in poorer deception ability.

Practice/Experience

The moderating factors listed above are largely immutable. That is to say, for example, that we cannot directly manipulate the aging process in order to mitigate multitasking ability. However, there are steps that can be taken that can improve multitasking ability. Research indicates that experience and practice can reduce dual-task interference (Ruthruff, Johnston, & Van Selst, 2001; Ruthruff, Van Selst, Johnston, & Remington, 2006). Beilok, Wierenga, and Carr (2002), for example, found that expertise doesn't require sustained attention and consequently frees up resources to allocate to secondary tasks. Moreover, Ericsson and Lehmann (1996) assert that experts can expand their working memory capacity to devote attentional resources to planning, reasoning, and evaluation. The benefits associated with experience and practice are principally a result of the automatization of individual tasks, which leads to reductions in the cognitive resources used during task performance (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Driving provides a good example of how this process occurs. When we first begin driving, we are challenged with the requirement to seamlessly steer, brake, check mirrors, monitor traffic, and so on. However, after years of practice these tasks become second nature. An "expert liar," for example, may be able to lie using only minimal attentional resources. Consequently, the "expert liar" has resource reserves that can be used for performing other concurrent tasks (e.g., monitoring the interviewer for feedback or controlling behavior).

Proposition 11: Experience and practice with lying can decrease the difficulty associated with telling a lie. As a result, individuals with greater experience/ practice may be more adept at lying and more difficult to detect. Additionally, research demonstrates that well-learned tasks impart a sense of control and predictability in the task environment (Keinan & Friedland, 1996). Perceptions of control and predictability may also benefit the deceiver.

Caveat: It is likely that the effect of experience and practice on enhancing deception ability may be limited to the types of lies and specific situations that are "practiced." That is, the advantages offered by increased experience or practice are likely to be context-specific. For example, we wouldn't necessarily expect experience with telling "white lies" to transfer to "high-stakes lies."

Training

Related to practice and experience, training can also improve multitasking ability (Bherer et al., 2005; Kramer, Larish, & Strayer, 1995). Training is a pedagogical approach aimed at enabling the acquisition of knowledge, skills, and attitudes (KSAs) through instructional methods (e.g., information,

demonstration, practice). A key distinction between experience/practice and training is the focus on KSAs. For example, skills such as time-sharing and attentional focus are certainly relevant components of multitasking. Targeted training interventions may focus on these specific skills (e.g., Heggestad, Carpenter, O'Shea, DeLosh & Clegg, 2002; Singer, Cauraugh, Murphey, Chen, & Lidor, 1991). Additionally, one aim of training is transfer of skills learned into a novel task environment. The underlying assumption is that specific skills should be generalizable across task environments and contexts. For example, the ability to time-share should benefit a new driver and a deceiver similarly. Previous research has shown that multitasking training can transfer to novel tasks (Bhemer et al., 2005).

Proposition 12: Individuals can be trained to become better at carrying out simultaneous tasks and, as a consequence, can become better liars. The published literature on what constitutes a good liar is almost nonexistent (see Vrij, Granhag, et al., 2004), and the published literature on how to train a liar is altogether nonexistent. Although these areas are neglected, this is to be expected (at least in the academic arena) because getting these types of studies past an institutional review board (IRB) may be challenging. It is not difficult to envision the reaction of an IRB reviewer when he/she recognizes that the purpose of a proposed study is to train participants to become better deceivers. Nevertheless, this type of research could be valuable. For example, this research could benefit law enforcement and military personnel who engage in clandestine operations.

Lie Strategy

Liars use different strategies when they lie and some of the strategies are less demanding than others. Although we are unaware of a taxonomy of deception strategies, we suggest that liars adopt one of three general strategies. The first strategy is an all-out lie that entails a complete fabrication. An all-out lie is the most demanding of the strategies and consequently the easiest to catch. The second strategy is to tell a simple lie. A simple lie requires less cognitive resources to manage and thus the liar is less likely to contradict themselves. However, the fact that a liar delivers a simple lie may act as an easily identifiable red flag. The final strategy involves liars delivering a lie about something they have done or have intricate knowledge of. This strategy is sometimes referred to as a half-truth. This type of strategy can be particularly successful because liars can tell a detailed story that doesn't require a great deal of cognitive resources to complete. For example, if an officer asks an individual about the activities they were engaged in on a specific date and time, they could easily provide a detailed description of the activities they did on the prior day.

Proposition 13: Cues to cognitive load and deception ability depend on the strategy the liar adopts. Conceivably, half-truth and simple lie strategies would be more difficult to detect than an all-out lie strategy.

Curvilinear Effect of Cognitive Demand

The effects of cognitive demand on performance are curvilinear. As Bowers, Braun, and Morgan (1997) state "it is widely accepted that the relationship between workload and individual performance is characterized by a curvilinear function where performance degrades at low and high levels of workload" (p. 87). This also implies that there is a level of workload that facilitates optimal performance. Beyond that optimal performance level, performance suffers and does so in an exponential fashion. In Fig. 8.1, the graph on the left represents a scenario in which Person A is a truthful person in conversation with an interviewer and Person B is a deceptive person engaged in fabricating a lie, monitoring the interviewer, suppressing the truth, and resolving discrepancies. In this scenario, Person A is not likely to demonstrate a performance decrement, nor the behavioral cues associated with cognitive demand. On the other hand, we would expect Person B to exhibit (to a degree) a performance decrement and the associated behavioral cues related to greater cognitive demand.

However, in this scenario, the demands of the task may not be sufficient to yield observable behavioral differences between Person A and Person B. As Vrij et al. (2011) suggest, certain cues may only become diagnostic when cognitive load is increased by deliberately imposing additional task demands. This scenario is represented by the graph on the right of Fig. 8.1. Through the imposition of additional task demands, Person B is strategically "pushed" into a zone of greater demand and correspondingly poorer performance; whereas we would expect the decrement to Person A's performance to be less evident. As a result, cues to deception are likely to be more diagnostic in this latter scenario.

Proposition 14: The fact that the cognitive demands associated with lying are greater than the demands stemming from telling the truth does not ensure that cues to cognitive demand will be more salient in any given situation. For example, under moderate levels of workload, a deceiver's task may not impose enough of a demand to impair his or her performance. However, as higher levels of
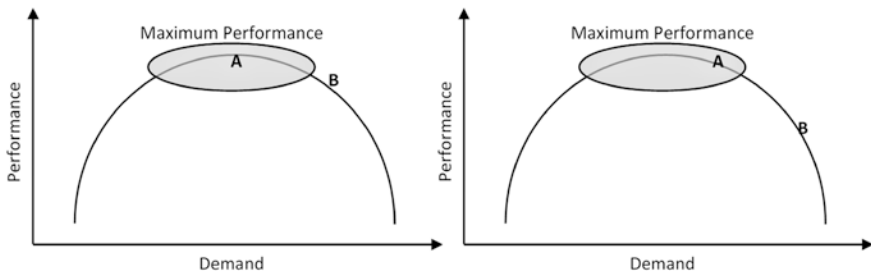
Fig. 8.1 The effects of additional cognitive load on discriminating truthtellers' from liars' performance

workload are imposed, performance is increasingly impaired. Because the effects of cognitive demand are curvilinear, cue salience will increase when the cognitive divide between truthtellers and liars increases.

DISCUSSION

This chapter sets out to link the cognitive load theory of deception to existing theories within the broader cognitive load literature. We did this not to expose a weakness in the existing literature (on the contrary, we acknowledge the significant impact the cognitive load theory of deception has had on the field), but to provide further validation for this approach and to identify research that the deception field can leverage to advance current understanding. The cognitive load approach to the detection of deception is supported by comprehensive theoretical approaches to multitasking and cognitive load. These theories are well supported, and although there are various perspectives offered by different researchers, the basic concepts and principles are clear, unambiguous, and offer a valuable lodestar or reference for future deception research.

We presented a number of propositions for future research throughout this chapter. However, caution should be taken before a specific avenue is comprehensively investigated. The effects of cognitive demand may seem straightforward, but they are dependent on a myriad of factors. As a result, researchers must be diligent in developing experimental protocols and recognizing the potential effects experimental manipulations and other factors may have on performance (e.g., effects may be additive or they may cancel each other out). For instance, we would expect the requirement to tell multiple lies during an interview to be more demanding than telling a comparable single lie. However, a performance decrement may not be demonstrated if the liar employs a deception strategy that mitigates cognitive demand (e.g., a half-truth). Accordingly, deception strategies should be assessed and controlled for. As another example, the requirement to tell multiple lies while, for instance, maintaining eye contact would likely have an additive effect. That is, it would be more difficult to do both than to do either separately. In short, it is the role of the researcher to play out the conceivable scenarios in order to mitigate unintended consequences.

Although the cognitive load approach to deception is supported by the existing theories of multitasking and cognitive load, the emotion-based approach to deception shouldn't be dismissed. On the contrary, the cognitive- and emotion-based approaches to deception are not mutually exclusive. In fact, parts of the emotion-based approach fit rather nicely into a cognitive outlook on deception. For example, the research on performance pressure (Baumeister, 1984; Hardy, Mullen, & Jones, 1996) outlines the effect emotions can have on cognitive functioning. Specifically, emotional responses and emotion regulation vie for the same limited attentional resources as

primary task performance. Consequently, the effects on cognitive demand and the resultant cues would be the same (i.e., a potential decrement on primary and/or secondary task performance). In this respect, the opposing approaches to deception work congruously with one another.

It is also important to note that the existing theories of cognition identified in this chapter relate primarily to sender communication (i.e., truthtellers and liars). Specifically, we focused on theories that could be used to explain variability in cognitive cues to deception. However, this only represents one side of the proverbial coin. The other side of the coin represents the receiver, or the individual attempting to diagnose deception. While the multitasking and cognitive load literatures would undoubtedly be applicable to lie catchers, there are a number of related fields that are equally pertinent. The classical and naturalistic decision-making literatures (e.g., Brunswik, 1955; Johnston, Driskell, & Salas, 1997; Kahneman & Klein, 2009; Kahneman & Tversky, 1979; Klein, Calderwood, & Clinton-Cirocco, 1986; Lipshitz, Klein, Orasanu, & Salas, 2001; Simon, 1955) and the dual-process literature (e.g., Chen & Chaiken, 1999; Evans, 2008; Petty & Wegener, 1999; Reinhard & Sporer, 2008) seem particularly relevant.

Concluding Remarks

Deception is a real and present danger to society. Developing a better fundamental understanding of the underlying mechanisms of deceit plays a vital role in mitigating these dangers. There is much to be learned from a more comprehensive understanding that stems from the existing literature on cognition and cognitive demand. The propositions advanced in this chapter provide valuable insight into how individuals may deceive and the factors that affect a deceiver's performance. These propositions are grounded in the broader literature on multitasking and cognitive load. As such, although they are presently untested, they offer potentially useful avenues to guide future research.

Notes

1. The term *workload* has several definitions. Generally, workload refers to an individual's perception of the work demands imposed by a task environment. The term also refers to the demands of the task environment itself in terms of the volume and pace of the work to be performed (see Young & Stanton, 2005).
2. It is important to note that multitasking is not the sole domain of liars. Truthtellers may be equally afflicted.
3. A related perspective on task demands is provided by research on *performance pressure* (Baumeister, 1984; Hardy, Mullen, & Jones, 1996). This line of research proposed that the pressure to perform well under conditions of high importance can paradoxically result in performance decrements. What has been termed *choking under pressure* has been explained by an increase in conscious

attention stemming from thoughts about the task, the situation, and its importance, thus reducing the resources available for task execution. Performance pressure represents a dual-task environment in which concerns about performance and task execution vie for limited resources (Beilock, Kulp, Holt, & Carr, 2004).

REFERENCES

Abe, N. (2011). How the brain shapes deception: An integrated review of the literature. *The Neuroscientist, 17,* 1–15. https://doi.org/10.1177/1073858410393359.

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London, 302,* 311–324.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Oxford University Press.

Baddeley, A. D. (1995). Working memory. In G. A. Gazzaniga (Ed.), *The cognitive neurosceinces* (pp. 755–764). Cambridge, MA: MIT Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). New York, NY: Cambridge University Press.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91,* 3–26. https://doi.org/10.1037//0033-2909.91.1.3.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46,* 610–620.

Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology: General, 133,* 584–600. https://doi.org/10.1037/0096-3445.133.4.584.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (2002). Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *The Quarterly Journal of Experimental Psychology, 55A,* 1211–1240.

Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., & Becic, E. (2005). Training effects on dual-task performance: Are there age-related differences in plasticity of attentional control? *Psychology and Aging, 20,* 695–709. https://doi.org/10.1037/0882-7974.20.4.695.

Bowers, C. A., Braun, C. C., & Morgan, B. B. (1997). Team workload: Its meaning and measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 85–108). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Broadbent, D. E. (1958). *Perception and communication.* Elmsford, NY: Pergamon Press.

Brookings, J. B. (1990). A confirmatory factor analytic study of time-sharing performance and cognitive abilities. *Intelligence, 14,* 43–59.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62,* 193–217.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories of social psychology* (pp. 73–96). New York, NY: Guilford Press.

Craik, K. J. (1948). Theory of the human operator in control systems: II. Man as an element in a control system. *British Journal of Psychology, 38,* 142–148.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

DePaulo, B. M., & Tang, J. (1994). Social anxiety and social judgment: The example of detecting deception. *Journal of Research in Personality, 28,* 142–153.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80–90.

Driskell, J. E., Mullen, B., Johnson, C., Hughes, S., & Batchelor, C. (1992). *Development of quantitative specifications for simulating the stress environment* (Report No. AL-TR-1991- 0109). Wright-Patterson AFB, OH: Armstrong Laboratory.

Driskell, J. E., & Salas, E. (Eds.). (1996). *Stress and human performance.* Mahwah, NJ: Lawrence Erlbaum.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*, 273–305.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59,* 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629.

Fischer, S. C., Morrin, K. A., & Joslyn, S. (2003). *Measuring multi-tasking ability.* Arlington, VA: Office of Naval Research.

Fisk, A. D., & Rogers, W. (2007). Attention goes home: Support for aging adults. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 157–169). Oxford, UK: Oxford University Press.

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience, 8,* 1298–1300.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Hancock, P. A., & Szalma, J. L. (Eds.). (2012). *Performance under stress.* Burlington, VT: Ashgate.

Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology, 87,* 621–636.

Heggestad, E. D., Carpenter, S., O'Shea, W. G., DeLosh, D. L., & Clegg, B. A. (2002). *Timesharing: Its future implications for the Navy.* Millington, TN: Navy Personnel Research, Studies, and Technology.

Huey, M. B., & Wickens, C. D. (1993). *Workload transition: Implication for individual and team performance.* Washington, DC: National Academy Press.

Johnston, J. H., Driskell, J. E., & Salas, E. (1997). Vigilant and hypervigilant decision making. *The Journal of Applied Psychology, 82,* 614–622.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D. (1975). Effort, recognition and recall in auditory attention. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance* (Vol. 6, pp. 65–80). New York, NY: Academic Press.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *The American Psychologist, 64,* 515–526. https://doi.org/10.1037/a0016755.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–292.

Keele, S. W. (1973). *Attention and human performance.* Pacific Palisades, CA: Goodyear Publishing Company.

Keinan, G., & Friedland, N. (1996). Training effective performance under stress: Queries, dilemmas, and possible solutions. In J. E. Driskell & E. Salas (Eds.), *Stress and human performance* (pp. 257–277). Mahwah, NJ: Lawrence Erlbaum.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fireground. In *Proceedings of the Human Factors and Ergonomics Society 30th Annual Meeting* (Vol. 1, pp. 576–580). Norwood, NJ: Ablex.

Kramer, A. F., Larish, J. F., & Strayer, D. L. (1995). Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied, 1,* 50–76.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science, 19,* 143–148. https://doi.org/10.1177/0963721410370295.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making, 14,* 331–352. https://doi.org/10.1002/bdm.381.

Mayr, U., Kliegl, R., & Krampe, R. T. (1996). Sequential and coordinative processing dynamics in figural transformations across the life span. *Cognition, 59*, 61–90.

Mcleod, P. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *Quarterly Journal of Experimental Psychology, 29,* 651–667.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. *Basic mechanisms. Psychological Review, 104,* 3–65.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica, 27,* 84–92.

Munsterberg, H. (1908). *On the witness stand.* Garden City, NY: Doubleday, Page & Company.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86,* 214–255. https://doi.org/10.1037//0033-295X.86.3.214.

Navon, D., & Miller, J. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology, 44,* 193–251. https://doi.org/10.1006/cogp.2001.0767.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually significant events. *Cognitive Psychology, 7,* 480–494.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 358–377. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6242412.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7972591.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories of social psychology* (pp. 41–72). New York, NY: Guilford Press.

Poposki, E. M., Oswald, F. L., & Chen, H. T. (2009). *Neuroticism negatively affects multitasking performance through state anxiety* (Technical Report for Navy Personnel Research, Studies, and Technology [NPRST-TN-09-3]). Millington, TN.

Reinhard, M.-A., & Sporer, S. L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology, 44,* 477–488. https://doi.org/10.1016/j.jesp.2007.07.012.

Riby, L. M., Perfect, T. J., & Stollery, B. T. (2004). The effects of age and task domain on dual task performance: A meta-analysis. *European Journal of Cognitive Psychology, 16,* 863–891. https://doi.org/10.1080/09541440340000402.

Ruthruff, E., Johnston, J. C., & Van Selst, M. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 3–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11248938.

Ruthruff, E., Van Selst, M., Johnston, J. C., & Remington, R. (2006). How does practice reduce dual-task interference: Integration, automatization, or just stage-shortening? *Psychological Research, 70,* 125–142. https://doi.org/10.1007/s00426-004-0192-7.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review, 115,* 101–130. https://doi.org/10.1037/0033-295X.115.1.101.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Shallice, T., Mcleod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? *The Quarterly Journal of Experimental Psychology, 37,* 507–532.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69,* 99–118.

Singer, R. N., Cauraugh, J. H., Murphey, M., Chen, D., & Lidor, R. (1991). Attentional control, distractors, and motor performance. *Human Performance, 4,* 55–69.

Smith, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin, 67,* 202–213. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5341896.

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14,* 1–36.

ten Brinke, L., Porter, S., & Baker, A. (2012). Darwin the detective: Observable facial muscle contractions revel emotional high-stakes lies. *Evolution and Human Behavior, 33,* 411–416.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging, 18,* 443–460. https://doi.org/10.1037/0882-7974.18.3.443.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities.* Chichester, UK: Wiley.

Vrij, A. (2015). A cognitive approach to lie detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: Current challenges and new approaches* (pp. 205–229). Chichester, UK: Wiley.

Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2002). Will the truth come out? The effect of deception, age, status, coaching, and social skills on CBCA scores. *Law and Human Behavior, 26*, 261–283. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12061619.

Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and reality monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioural Science, 36,* 113–126. https://doi.org/10.1037/h0087222.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22,* 1–21.

Vrij, A., Fisher, R. P., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling, 5,* 39–43. https://doi.org/10.1002/jip.

Vrij, A., Granhag, P. A., & Mann, S. (2004). Good liars: Our invited article in *Review of Policy Research* that never got published (May 2005), 1–20.

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science, 20,* 28–32. https://doi.org/10.1177/0963721410391245.

Vrij, A., & Mann, S. (2001). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15,* 187–203.

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32,* 253–265. https://doi.org/10.1007/s10979-007-9103-y.

Vrij, A., Mann, S., Leal, S., & Fisher, R. P. (2010). "Look into my eyes": Can an instruction to maintain eye contact facilitate lie detection? *Psychology, Crime & Law, 16,* 327–348. https://doi.org/10.1080/10683160902740633.

Welford, A. T. (1967). Single channel operation in the brain. *Acta Psychologica, 27,* 5–22.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 63–101). New York, NY: Academic Press.

Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple task performance* (pp. 3–34). Bristol, PA: Taylor & Francis.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3,* 159–177.

Wickens, C. D. (2005). Multiple resource time sharing models. In N. A. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Human factors and ergonomics methods* (pp. 40.1–40.7). Boca Raton, FL: CRC Press.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50,* 449–455. https://doi.org/10.1518/001872008X288394.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2016). *Engineering psychology and human performance* (4th ed.). New York, NY: Routledge.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 30,* 599–616.

Wright, G. R. T., Berry, C. J., & Bird, G. (2012). "You can't kid a kidder": Association between production and detection of deception in an interactive deception task. *Frontiers in Human Neuroscience, 6,* 1–7. https://doi.org/10.3389/fnhum.2012.00087.

Young, M. S., & Stanton, N. A. (2005). Mental workload. In N. A. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Handbook of human factors and ergonomics methods*, Ch. 39. Boca Raton, FL: CRC Press.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). New York: Academic Press.

Discursive Dimensions of Deceptive Communication: A Framework for Practical Analysis

John H. Powers

Using language to communicate our ideas with others is among the most distinctively human things we do. Through language, we can mention important things in our experience, share information we believe to be true, cooperate to accomplish our goals, describe the things we see or imagine, offer taxonomies of items in a set to reveal how they interconnect, tell stories that hold some group of events together as a coherent whole, and provide reasons to believe or act in a certain way. However, for all of the wonderful things language may be used to accomplish, it can also be used as a major instrument of deception during interpersonal, group, and public communication in order to lead others to draw wrong conclusions based on what someone has said.

Deception is different, of course, from mere error; *discursive deception* is the intentional use of language to mislead, misdirect, or misinform another person in order to induce them to follow a flawed path of thinking, belief, or behavior. When claims are made, for example, about the size of a crowd or the degree of voter fraud during an election when there is strong evidence to the contrary (e.g., photographs taken at the scene; reports received from the voting commissions in each state), the person making the claim could simply be in error or could be intentionally trying to mislead the audience receiving

J. H. Powers (✉)
Professor Emeritus of Communication Studies,
Hong Kong Baptist University, Kowloon Tong, Hong Kong
e-mail: jpowers@associate.hkbu.edu.hk

the message. In this chapter, I analyze how language can be consciously used to mislead others when the sender of the message has reason to want to distort the truth for their personal benefit. However, because a sender's *intent* to deceive is perilously difficult to prove, much of the discussion below can also be applied to understanding a sender's linguistically induced error as well as self-conscious deception.

The underlying theme of this chapter is that "words matter." It will describe a framework of concepts and principles for analyzing deceptive discursive practices that can occur at four different levels of language, moving from the smaller units of discursive communication to the larger ones. Specifically, the first section looks at deception arising at the lexical (i.e., individual word choice) level of discourse; the second explores the propositional or syntactic bases of deception; the third section surveys the speech act dimensions of deceptive discursive practices; and the fourth features the types of deception that can arise from the macrosemantic dimensions of discursive communication such as descriptions, narratives, and argument structures. Each section will introduce a small number of principles related to a particular level as a starting point for identifying the modes of deception that are made possible by abusing the key principles of language at that level.

ANALYZING THE LEXICAL/SEMANTIC FOUNDATIONS FOR DISCURSIVE DECEPTION

The lexical/semantic level of discursive communication focuses on the substantive words that are chosen to name the things that a communicator is talking about. Generally speaking, these are the nouns, verbs, adjectives, and adverbs that make up the vast majority of the words in any language. Almost all linguistic deception begins with a *contest* to label something in a particular way in order to produce a desired result from the person being deceived. To see why this contest over the "right" label for talking about something is fundamental to understanding many forms of discursive deception, we can summarize some basic principles related to the lexical level of language.

First, these types of words always identify *categories* of things into which many items may fit, even when they are used to mention specific instances of the category for a particular purpose at a particular time. In a sentence like "My computer is broken," we have two category-naming words: "computer" and "broken." The words are used to identify the categories that the speaker wants to mention, and the sentence applies them to talk about a specific instance of the category, namely "my specific computer" and "its current condition."

Second, every category-naming term may be defined by a set of features that characterize the category—i.e., by a set of characteristics that something needs to display in order to be a member of the category. When we look up a word in the dictionary, at the very minimum the entry provides a list of such

features or attributes. If we know the meaning of a word (such as *computer*), we generally know the feature set that is relevant to understanding what is being said. Therefore, to call something "X" is to imply that, at the minimum, it manifests the features that define the category named by the word chosen.

Third, the category labels chosen on a particular occasion have consequences. When we select a particular category labeling word instead of another to talk about an item, we are inherently highlighting the feature set we want to treat as the most salient for the discussion. *Select a different label to mention whatever we are talking about and we invoke a different set of features to think about it* because the category labeling term chosen emphasizes certain features or details about the object and hides others that would be emphasized if a different category-naming term had been chosen.

Fourth, this means that the category labeling term used to talk about something inherently "tunes" the mind of the other in a way that, however transiently, aligns the receiver's thinking with that of the sender's (Brown, 1958/1970). Why would one want to do this? Most likely, it would be to lead the audience to consider the consequences for action that the sender most prefers to be the outcome of a discussion. For example, if someone like Edward Snowden is labeled as a *spy* rather than a *whistleblower*, *traitor*, or even as an *informant*, there are different legal consequences of the category label chosen. Similarly, are countries such as China or Russia best categorized as *enemies, adversaries, competitors, antagonists, opponents, rivals*, or *foes*, and so forth? We would need to know which features differentiate these terms from one another as part of determining what the consequences for action would be of choosing one of these terms rather than another.

With this introduction to the principles of category construction, we can now apply the principle of categorical labels and their mind-tuning feature sets to understanding several different types of lexical deception, beginning with the choice to use a misleading category labeling term in the first place.

Deception by Offering Misleading Categorical Labels

Categorical deception occurs when a sender knowingly selects a category label whose attribute set is not the best fit for the object referred to but which is more likely to favor the outcome the sender most prefers. Thus, for example, when the US Secretary for Housing and Urban Development, Dr. Ben Carson, referred to African slaves as *immigrants* to America, he ignored critical features that distinguish the category *slave* from that of *immigrant*. For instance, slaves were brought to the US involuntarily, while immigrants exhibit the feature of voluntary choice. To begin to think of slaves as just a subcategory of immigrants is to fail to think about the types of brutality involved during the enslaving process and the social consequences for the generations that followed. These two terms have quite different defining

feature sets and also have quite different consequences for how one treats the person who is being labeled. What makes such words important for verbal deception theory is that one feature set may more accurately reveal what is relevant to a particular discussion and another one effectively hides some of the salient features that apply to developing a national housing policy. Below are additional examples of competing categorical labels that would lead to vastly different consequences if one category label rather than another becomes the preferred way of talking about the categorized object.

- When is an action a "hate crime," "terrorism," or simply "murder"?
- What is the difference between a *refugee* and an *economic migrant*?
- Which category is most suitable for talking about persons coming across the southern border of the US without a valid visa: *illegal immigrants* or *undocumented workers*?

In each case, we can ask why one side of the discussion would want a particular label to be used while another would prefer the alternative term?

Deception can occur in these cases when a communicator chooses a categorical label that allows the person to pursue a particular goal because the category chosen would itself normally be treated in a particular way. For example, if the government wants to deport all undocumented foreign nationals who have committed a *crime*, classifying Driving Under the Influence (DUI) as a "crime" makes that goal much easier to pursue than classifying it as a "misdemeanor." An even more stringent application of the manipulation of categories to accomplish a particular goal is classifying the mere fact of being "undocumented" as itself a *crime*, thereby extending the potential for immediate deportation to a vastly larger population of "eligibles."

We may now explore some additional modes of lexical deception that arise from the "mental tuning" effect of the categorical labels one chooses.

Deception by Proposing Misleading Contrastive Labels

There is a sense in which every word choice we make when talking about something is "contrastive" with every other word we could have chosen but didn't. The examples above all illustrate some contrastive options available to a deceptive communicator.

However, *contrastive* deception means something slightly more specific than just selecting one's most preferred label and using it. **Contrastive deception** occurs when the sender acknowledges that an alternative label *could* be used but then explicitly rejects that category in favor of a preferred (but potentially misleading) label so that the dispreferred alternative is treated as almost morally unacceptable for thinking about the topic. Thus, for example, when a Muslim teenager was bludgeoned to death with a baseball bat

while on the way to a late-night Ramadan event with her friends, the police immediately classified the killing as "road rage," and publically dismissed the possibility of classifying it as a "hate crime," even though the girl was wearing a traditional Muslim head scarf and was walking toward a nearby mosque for a religious event (Magane, June 21, 2017). Whether rightly or wrongly, the police explicitly rejected the possible "hate crime" elements of the circumstances and considered only the far less severe category of "road rage"—a category that would make the killer less "responsible" for the crime and which would also result in a much different punishment at trial.

Deception by Emphasizing Agonistic Oppositions

To be in an agonistic relationship is to be in conflict with some opposing person or force. In language, agonistic oppositions are contrastive words (usually in bipolar opposition) that are set up to highlight conflicting labels for naming something, with one member usually being treated as having a positive valence and the opposing term being given a negative valence. This results in potential deception because an either-or type of villainization is artificially set up around the agonistic opposition. For example, a speaker might say in regard to a suicide bombing that the world is now engaged in a war between *barbarism* and *civilization*. Here, the speaker assumes that there are two clearly agonistic "sides" and that the suicide bomber represents *barbarism* and the affected community is the exemplar of *civilization*.

Deception by Shifting Abstraction Levels

Categories come into being as the result of the mental process of abstraction, wherein the category necessarily ignores the particular details of different items in order to focus only on a set of features that are common to all members of the category. "My computer" may be different from yours in that mine is a laptop model, and yours is a desktop model. My monitor may be 15 inches diagonally and yours 27 inches; mine may be silver gray and yours jet-black. However, when we mention "my computer," all of these differentiating details are ignored through the process of abstraction because all computers share a common set of features (e.g., the ability to use software to process and display information) that we recognize as defining the category "computer."

Because of the nature of the abstraction process that makes category formation possible, category labeling terms on most topics can be arranged into hierarchical "ladders" (Hayakawa & Hayakawa, 1990), with each higher category on the ladder being more abstract than the category at the next lower rung. The more abstract categories have fewer limiting features. Therefore, two words might both be used to name a referent in the category but require a different number of features to become a member of the category. Thus, for

example, we could arrange categorical labels like these, from most abstract to least:

<div align="center">

Object
Machine
Computer
Laptop

</div>

These terms are contrastive, but not because they have different *types* of features. Rather, the contrast arises because they have different *numbers* of features and therefore different levels of abstraction. A laptop (the most specific set of features) is a type of computer, a computer is a type of machine, and a machine is a type of object (the least specific set of features). Thus, the category "object" is far more abstract than is the category "laptop." The four terms are displayed as a hierarchy of related but contrastive terms based on their degree of abstraction.

The ability to place categories in an "abstraction ladder" series opens considerable room for deceptive verbal practices. This can occur because using the more abstract term can lead one person to think of one instance of a subordinate category and another person to conceive a different, sometimes quite incompatible, subordinate category. Yet they can both think they understand one another even if their cognitive tuning has become quite distant from one another.

A clear example of the deceptive potential that different categorical levels on an abstraction ladder produce is *euphemism*, which frequently uses a more highly abstracted label to hide unpleasant details that would necessarily emerge by using a category word from a lower level of abstraction. For example, in April 2017, when United Airlines forcibly removed a seated passenger to make room for a flight attendant who needed to be available for a flight in another city, United reported that it had "reaccommodated" all passengers who had "volunteered" to leave the plane (Steinmetz, April 12, 2017). Some of the passengers might have felt that they had been "voluntarily reaccommodated," but it seems unlikely that the forcibly removed passenger, who was bleeding profusely and who lost two teeth in the violent episode, would have thought United's abstract category label properly applied to his situation.

Deception by Using Words with Vague Boundary Conditions

The boundary conditions of every category can become uncertain at their edges, where it may be hard to determine whether or not a particular defining criterion fits a situation or not. But there are also categorical words that are inherently fuzzy—where we do not know which defining criteria might apply in deciding if the category applies. Under these conditions, we tune our own minds as best we can and assume we have understood the speaker's intent because we do not know which defining criteria the speaker wants to apply.

When using words that are inherently imprecise concerning which criteria that items must exhibit in order to be members of the category, deception can be carried out via vague categorical labels. For example, if a school or, as happened in 2017, an airline insists that its passengers must "dress modestly" while riding with that airline, is it possible to be certain what is meant by this policy-labeling term? Or can the purveyor of this policy arbitrarily decide that a particular clothing choice on a particular occasion is "immodest" and disallow the individual from flying based on a vague and arbitrarily applied category?

Deception by Using Misleading Metaphors

Metaphor is the process of using the features that define a category labeling term drawn from one area of our knowledge (the source) to try to gain insights into the features of some other area of experience (the target) that seems similar in some way. We use metaphor to point to a perceived similarity between the features of the source concept and the target area of experience that we might not otherwise see. Thus, we might say that Y (target) is like X (source) in sharing the following defining features (Lakoff & Johnson, 1980). For example, if someone says that "students are the customers of the university," the speaker is trying to provide insights into the student–teacher relationship by drawing insights from a different realm of activity, namely the seller and buyer relationship between a merchant and a patron.

During the metaphoric process, there is a large potential for both error and deception. This is possible because metaphors cannot highlight only full similarities; they can also hide dissimilarities or distort partial similarities to make them seem like full similarities. For example, the person who thinks students are like customers may emphasize that the teacher is a supplier of something valued and that the student's tuition represents the purchase of the thing valued. Given this interpretation, students may feel that they are owed a grade or that they must be allowed to try again to redo a failed assignment, just as they would be guaranteed by a seller to be able to "use" a working product that they had bought. All of this begins with the metaphorical starting point that schools charge tuition for a course and the student pays tuition for the class. For those who reject this metaphor as deceptive, they might point out that it fails to reveal other aspects of the teacher–student relationship, or what it is that tuition actually pays for. Indeed, a teacher might find that the purchase price (i.e., tuition) is more like buying a chance to try one's luck or skill in a contest. Not everyone succeeds just because they put their money down.

Deception by Using Misguided Metonymies

Metonymy is the discursive practice of using the name for something that is merely *associated* with a person, process, or object as the temporary name for that person, process, or object. The normal goal of metonymy is to highlight

a feature of the thing being named that is particularly salient to the sender on a particular occasion. If I say "can you give me a hand here" when I am asking for help, the categorical word *hand* stands in metonymically for the word *help* in order to emphasize the kind of help I want. We see metonymy in a sentence like "The Whitehouse said today that …" Why do this when we could say, "a spokesperson for the president said …"? The reason for choosing to identify the "spokesperson" as "the Whitehouse" is to indicate level of importance of the speaker and the message. In the current era, one might report that "Mar-a-Lago said …" But if we did, what features would we be highlighting by this shift of metonymic reference?

For the purpose of understanding deception, we may identify two types of misguided metonymies, which we may label as *mistaken metonymies* and *metonymic exaggerations*. **Mistaken metonymies** are created when the associated category chosen for the metonymy does not actually apply. That is, the metonymy is imposed on the item being named rather than being a naturally available association. "Shorty" for a very tall man, for example. **Metonymic exaggerations** are created when a characteristic that is actually associated with the object (but only weakly) is made to appear as more prominent than it really is. During the presidential primary election season of 2016, candidate Trump created both exaggerated and mistaken metonymies for each of his primary opponents. By constant repetition of such phrases as "Little Marco," "Lyin' Ted," "Low Energy Jeb," "Crooked Hillary," and even "Pocahontas" for non-candidate Elizabeth Warren, Mr. Trump was trying to establish a seemingly natural metonymic relationship between the descriptive term he used and the opposing person he named.

ANALYZING THE PROPOSITIONAL/SYNTACTIC FOUNDATIONS OF DISCURSIVE DECEPTION

Individual category-naming words are not themselves deceptive until they are used in sentences to make statements. Categories were given considerable attention in the previous section because of their critical importance once they have been selected for use in sentences in order to assert claims. The difference between a list of categorical words such as *green*, *moon*, and *cheese* and a sentence such as "the moon is made of green cheese" makes this obvious. Thus, the second level of language we need to explore concerns the types of deception that are possible when we are uttering simple declarative sentences that may be deceptive for different reasons.

To get our discussion started, it is useful to introduce two concepts: *proposition* and the process of *asserting* a proposition in order to make a claim. *Propositions* are word pictures of a state of affairs; a proposition is constructed by selecting category labeling words to name the elements that make up the picture expressed: moon + made + green + cheese collectively describe a state of affairs. *Assertion* is the process of saying that the proposition is true: "The moon is made of green cheese." The assertion process can claim that

the word picture is true now, was true in the past, or will become true in the future. In this case, however, we know the assertion of this particular proposition is not true and, if we seriously tried to pass it off as true, it would be a deceptive statement.

Next, we need to break the process of assertion into its two key parts: reference and predication. The *reference* in a sentence is what is being talked about; for this reason, the referent is sometimes called the "subject" of the sentence. In our example, the referent is "the moon." The *predication* of a sentence is everything that is said about the reference. Generally, in English sentence structures, the referent is everything in a sentence that comes before the primary verb; the predication is everything said about the referent, starting with and including the primary verb. Accordingly, "is made of green cheese" is the predication. Now let's apply this to understanding deception at the sentence level of the assertion.

Deception by Making False or Misleading Predications

At the propositional level, the most obvious aspect of deception is the making of a truth claim that is false in its entirety (i.e., the *bald-face lie*). Bald-faced lying typically occurs when one knowingly predicates false information about the referent of a sentence so that the assertion as a whole is false. Thus, if someone says that a New York pizza shop (the referent) is a secret place for illegal activities (the predication) when it is really just an ordinary pizza establishment, the statement is a bald-faced lie of this type.

Deception by Mentioning Misleading Referents

Although lying is usually accomplished by stating something in the predication of an assertion that is untrue of the referent, it is also possible to be deceptive by mentioning a referent that either does not exist as if it *did* exist or by identifying the wrong referent required to make the claim true. The deception is enacted in the referential part of the sentence rather than in the predication. The classic example is found in the sentence, "The king of France is bald." Typically, we would look at the truth value of the predication to determine if the king has a full head of hair or not. However, the deception/error is not in the predication about the king. It is in the referent "the king of France" because there is no longer a king in France. Thus, the deception rests in using a referring term that mentions something that doesn't exist but which is falsely treated as if it does. Similarly, mentioning "the war on coal" as the referent in a sentence such as "The war on coal is dead" fits exactly this pattern of deception/error. It mentions a referent that does not currently exist as if it did and then it predicates something about it. The deceived person may then search for the truth or falsity of the predication without realizing that the referent itself is where the problem arises because the referent does not actually exist. To expose the deception,

one might ask the sender if there really has been a recent "war" on coal, or is it just the case that the coal industry is in the natural process of transitioning toward irrelevance in the face of the new modes of energy production that are being invented? This is the same kind of question as "Was there ever a *war* on horse drawn carriage makers self-consciously aimed a putting them out of business?" Or did the rise of the horseless carriage simply supplant the need for a large number of horse-drawn vehicles?

Deception by Using Embedding Syntactically Hidden Propositions

Communicators may use grammatical principles to efficiently express multiple propositions in a single sentence. If I say "My brother in Phoenix just got a new job," the sentence expresses three propositions, one visible and two that are relatively invisible. The obvious proposition is clearly stated: "My brother got a new job." The two relatively invisible propositions are:

> I have a brother.
> My brother lives in Phoenix.

These were made invisible as separate propositions through the magic of English syntax. But they are equally asserted by the sentence as was the main proposition. We just don't experience these assertions as directly because they are cleverly embedded in the reference portion of the sentence.

Why is this important for understanding the grammatical potential for deception? Perhaps it is obvious, but any proposition embedded in a sentence through grammatical means could be either true or false, just as the main proposition explicitly expressed in the sentence can be either true or false. Accordingly, deception may arise from the claim that is actually stated "visibly" or via another, non-visible means. Visible falsehood is what we usually mean by a lie. The person asserts an overt claim that he knows to be false. The claim being made is visible to both of the communicators, though only the sender will know that it is false. To appreciate the potential for deception in invisible propositions, suppose I had said, "My sister in Kansas City just got married." The invisible propositions that are presumed by the receiver to be true are that I have a sister (I don't) and that my non-existent sister lives in Kansas City. The deception is that I have a sister and that sister lives in Kansas City, though I never directly asserted either of those things overtly. I only said she was recently married.

Deception by Imposing Invisible Presuppositions

Presuppositions may be defined as any unstated proposition that must be true before the stated proposition can be true. That is, they are the hidden assumptions that are presented by the deceiver as already accepted as true

before the overtly stated claim is made. They are different from the syntactically hidden propositions described above because they don't need to be part of the syntax of the assertion in any obvious way. So, while it would be true to say that the sentences above presuppose that I have a brother and a sister, here we are looking more deeply into what the deceiver is assuming to be true and hopes the receiver of the message will also accept as true without the need for arguments or evidence.

We play with presuppositions of this type when we tease, as in the question, "Have you stopped beating your partner?" The question presupposes not just that the person has a partner relationship with someone, but more sinisterly that we know or believe that they have a violent relationship with them as well. So let us see how this type of deeply hidden presupposition might be used in a realistic context.

After the war in Afghanistan began in 2001, President George W. Bush explained his reason for bombing there, with words to the effect that "America's actions in Afghanistan are punishment for evildoers." What are the presuppositions that the president was asking the American people to accept without further explanation or evidence? Here are at least a few:

Presupposition : Evil has been done.
Presupposition : America knows who is responsible for the evil.
Presupposition : America has the right to punish without a trial those who are responsible for the evil.
Presupposition : Bombing Afghanistan is the appropriate punishment for the evil that has been done.
Presupposition : The people who did the evil will be the only ones punished for their actions by the bombing.

The key here is to recognize that these presuppositions are not directly retrievable from what was actually said, yet they are the hidden foundations that are meant to be accepted based on what *was* said. For this reason, they can be profoundly important for understanding the potential for deception when people speak. In the public sphere, these are the kind of presuppositions that reporters would need to probe deeply in order to fully illuminate what politicians on any side of a political issue are actually advocating.

Deception by Proposing Spurious Entailments

When we assert the truth of a proposition, there are usually consequences that follow from the fact that the proposition is accepted as true. Consequences that directly follow from the truth of a claim are generally known as *entailments*. Entailments are invisible propositions that are assumed to truthfully follow if the overt truth claim is true. These, then, are the unstated consequences that someone must also accept as true by adopting the deceiver's explicit proposition.

Entailments may be logically true or socially/culturally true. Logical entailments are propositions that must be also true if the overt claim is accepted as true, i.e., what follows from the words used to make the truth claim. If I say that "Jack got married recently," the fact that "Jack is no longer a bachelor" is logically entailed (guaranteed to be true) simply by meaning of the word *bachelor*, i.e., someone who is married is not a bachelor.

In addition to the way that word meanings are logically connected to allow logical entailments, social customs and cultural understandings can also create what we might call social entailments. And it is here that there is considerable space for using our language deceptively. *Social entailments* are those consequences that we accept as true based on social knowledge, beliefs, customs, prejudices, and so forth, about things that are outside of language proper. In these, the receiver of the visible truth claim is expected to supplement the truth claim by assuming that a certain entailment also follows because that would be the normal, socially accepted consequence of the claim being true. Thus, for example, to assert that "Mr. X is a billionaire" might entail for some people that "Mr. X is unconcerned with the social misfortunes of those in the middle class" because that is a common cultural association attaching to the concept of "billionaire." Thus, to mention Mr. X's economic status could be a deceptive entailment when talking about Mr. X in any other context. Indeed, during the 2016 US presidential primary season the phrase "the billionaire class" was frequently used with precisely this social entailment.

Deception by Imposing the Criteria for Judging One Type of Claim on Another

There are several different types of claims that someone might assert to be true, but each one has a different set of criteria that need to be applied in order to determine the acceptability or deceptiveness of the claim proposed. Five common categories of proposition are fact, value, policy, definition, and classification claims. Only one of these is properly subject to the criterion of being true or false (the factual claim), but all of the others can be uttered in such a way that they sound as if they are truth claims and therefore can result in a deceptive act. One of the most obvious is treating a value claim as if it were a factual claim: *Gone with the Wind* is a better film than *The Wizard of Oz*. The value foundation of this claim is signaled by the evaluative word "better," but the sender passes off the claim as a factual one subject to the criteria for evaluating truth claims rather than those attaching to value claims. Similarly, the communicator who says that the government should do X is proposing a policy for future action not stating a factual claim. In either case, the speaker could argue on behalf of the value or policy claim, but this would need to follow different criteria for proof (Powers, 2016). Short of that, passing a value claim or policy claim off as a factual claim and therefore mistakenly using the criteria for evaluating the truth of the claim can be deceptive to the receiver.

Deception by Adopting Misleading Grammatical Structures

Finally, deceivers may use the grammatical resources of their language to organize information in their assertions in ways that highlight or hide features of the underlying propositional structure. For example, the deceiver can hide the *agency* of a claim (i.e., who is responsible for an action) by using the passive tense. Imagine a congressperson opining at a town hall meeting the following: "It was decided that covering preexisting conditions is too expensive and that they should not be included under the revised government supported health care plan." The constituents might want to ask *who* decided this so that they hold that person or group responsible for the action—especially if the congressperson standing before them is among those who voted in support of the "agentless" bill.

ANALYZING THE SPEECH ACT/PRAGMATIC FOUNDATIONS OF DISCURSIVE DECEPTION

In the book *How to Do Things with Words*, J. L. Austin (1962) argues that language can be used for accomplishing more things than merely asserting true or false propositions. Indeed, he explains that when we talk, we are usually *doing* something as well as *saying* something. For example, we may be promising, threatening, cajoling, and so forth, and none of those activities involves asserting a truth claim. Austin's work has subsequently fostered an entire discipline called pragmatics and has also inspired an approach to linguistic pragmatics generally called speech act theory.

At the speech act level of discourse, we can identify two aspects of language with the potential for deceptive use. The first arises from Austin's original theory of performative speech acts, especially as developed in John Searle's (1979) taxonomy of five types of speech acts. For one can deceive by appearing to be performing one type of speech act when actually intending another. A second approach to speech act deception is clearly suggested by H. P. Grice's (1975) theory of *conversational implicatures*. The theory aims to explain how uttering certain words can lead the receiver to wrongly draw a conclusion from what the deceiver says because the receiver will apply normal "conversational maxims" to interpret the deceiver's words. Accordingly, this section will briefly introduce each of these three theorist's major ideas and illustrate how those ideas may be used to understand deception at the speech act level of discourse.

Deception by Abusing the Normal Speech Act Rules

Austin's key idea is that learning to use language is like learning the rules of a game. To the extent that two people know how to use a language, they have internalized a set of game-like rules for accomplishing their goals (such as promising, threatening, naming, and so forth). As part of his theory,

Austin developed a set of general rules that people seem to follow when they are doing what are called "performative" speech acts. In parallel, he developed a set of terms for describing how one's speech act intentions could fail (or be abused) that provide a theoretical framework for analyzing the potential deceptive uses that can be made of the rules.

According to Austin, for any kind of speech act to be successfully performed, there must be (a) an accepted conventional procedure, (b) having a certain conventional effect, (c) accomplished by uttering certain words, (d) by certain persons, (e) in certain circumstances. These persons and circumstances must be (f) appropriate for the procedure invoked, and the procedure must be executed by all participants, (g) correctly, (h) completely, (i) sincerely, and (j) with the intention to comply in the future.

For each element of these rules, there is a corresponding way the procedure can fail, say, for example, by what Austin calls misfires and abuses—i.e., acting as if a particular speech act has been performed when for some reason it has not. And it is these ways that speech act intentions can fail that can also lead to a corresponding form of deception. For example, the deceiver can try to persuade the other (a) that there is a conventional procedure for accomplishing something when in fact there is not, (b) that the expected conventional effect has been accomplished when it has not, (c) that the proper words have been spoken when they have not, (d) that the proper persons have performed the speech act when they are not in fact such persons, (e) that all of this was done in the appropriate circumstances when they were not, (f) and that all of these things were performed, (g) correctly, (h) completely, (i) sincerely, and (j) that they were all done with the intention to comply in the future when they were not. In this system, we have identified at least 10 possible modes of linguistic deception that have nothing to do with whether or not the utterance is literally true or false.

Deception by Misusing Indirect Speech Act Types

Building on some of Austin's ideas, philosopher John Searle (1979) developed a taxonomy of five general categories of speech acts, distinguished based on how the words uttered connect to the world outside of language. What he calls *assertives* and *expressives* both use language to try to match the words uttered to the world as it is. More particularly, however, assertives point to the world outside the speaker who utters them. This is Searle's way of identifying the conventional truth claim made about situations in the world. In contrast, expressives point to the world inside the person uttering the words. They are marked by words such as "I wish," "I hope," "I think," and so forth. They get their name because these utterances *express* the person's feelings and ideas, but are not verifiable because they are statements about the person's psychological dimensions that are only available directly to that person. We can hope that persons who say "I love you" are sincerely reporting their feelings, but we cannot really know the truth of such self-reports with certainty.

What Searle calls *commissives* and *directives* are words uttered with the intention of having some aspect of the world subsequently come to match the words that are spoken. For example, when someone says "I promise to pay you Tuesday for a hamburger today," the speaker is making a commitment to influence some small aspect of the world in the future. A promise commits the person who is speaking to try to make the world match the words about sometime in the future. In contrast, directive speech acts try to induce another person to change the world in a specific way sometime in the future. Thus, the parent who says "Please clean up your room after you finish your homework" is trying to direct the other's future behavior in a specific way.

The final type of speech act in Searle's taxonomy is the *declarative*. Here, under the right conditions, the words immediately change the world the moment they are uttered. Thus, for example, if during a wedding ceremony the participants say "I do," followed by the words from the officiant, "I now pronounce you man and wife," they are married. Their world has changed.

Speech act theory also recognizes that there can be both directly explicit examples of these categories as well as indirectly implicit instances. Thus, if the parent above says "Your room sure could use some dusting," they are probably not simply making an assertive observation. Instead, they are more likely using a polite form of directive to move the person to clean the room in the near future. Similarly, if a person says politely, "Can you pass the salt," for the other to simply reply "yes" but not to do so would seem to be either rude or incompetent. The question yes-no form was intended as a directive, and not as an informational question simply to be answered.

Now we can connect Searle's taxonomy of speech act types to the potential for deception. As we have just seen, utterances of one type can be used to accomplish social goals of another type without needing to say them directly. Thus, deception can be accomplished by appearing to perform one type of speech act while actually intending to perform a different type. Among other things, this opens up the deceptive practice of "plausible deniability" because one can always say, "But I didn't actually say that." In the example of the child being indirectly instructed to clean the room by way of an assertive statement, the parent cannot be accused of having pronounced "an order" to a teenager who may not want to feel "bossed around" by the parent. But how does someone know that what was uttered is intended differently from the form of what was actually said? For that, we need an additional bit of theory, namely H. P. Grice's theory of conversational implicature.

Deception by Flouting the Cooperative Principle and Its Associated Maxims

During US Senate Intelligence Committee hearings on June 8, 2017 (*New York Times*, June 8, 2017), former FBI Director, James Comey, reported that President Donald Trump said to him, "[Michael Flynn] is a good guy and has been through a lot. I hope you can see your way clear to letting this go, to

letting Flynn go. He is a good guy. I hope you can let this go." Mr. Comey reported to the Committee that "I had understood the President to be requesting that we drop any investigation of Flynn." How is it that the utterance beginning with "I hope" could be understood by Mr. Comey as an indirect directive to drop an investigation into Flynn's behavior? Here, we have a clear example of an expressive *form* utterance ("I hope") being interpreted as a directive for future behavior. But how could Mr. Comey have arrived at his conclusion that the expressive form was really intended as a directive speech act?

Questions such as this led philosopher H. P. Grice (1975) to develop a theory of conversational discourse that emphasizes the cooperative behavior involved in interpreting such utterances. Grice's answer is that conversationalists are able to *reason* their way from the literal statement to its likely intended meaning based on a set of rational principles governing how conversations may be understood. These principles, in effect, serve as logical premises conversational participants use to start their reasoning process. Grice organizes his theory into three main parts: (a) the Cooperative Principle itself, (b) a set of Conversational Maxims that conversationalists use as the rational basis for cooperatively interpreting conversational contributions, and (c) the Conversational Implicatures that provide the basis for selecting rational interpretations of non-literal utterances.

Grice's **cooperative principle (CP)** asserts that all conversational participants necessarily presume that the other participants intend to cooperate in maintaining the conversation. In practice, the CP means that all parties will assume that anything someone says is intended to be a purposeful contribution to advancing the conversation. If something is said that is not transparently cooperative, the listener will try to find a reasonable interpretation of the utterance that would reveal it to be cooperative after all. The classic example of this occurs when a person asks a yes-no question and receives a reply such as "Is the Pope Catholic?" Because this response is not *transparently* answering the original question, the first person must try to figure out what is intended by what is said and usually realizes that the answer was a playful (but still cooperative) affirmative answer to the question. So, how is this done?

According to Grice, there are four *categories of cooperativeness* that are necessary for implementing the spirit of the general CP, which he calls quantity, quality, relation, and manner. They provide the logical principles people use as they reason their way to an adequate interpretation of what is meant by what was said when an utterance is not transparently cooperative. Under each category, Grice suggests a number of specific "maxims" or principles that people expect each other to follow.

For example, under the category of **quantity**, Grice suggests two maxims that guide participants in determining *how much* to say during a given speaking turn.

1. Make your contribution as informative as is required (for the current purpose of the exchange).
2. Do not make your contribution more informative than is required.

For **quality**, Grice again proposes two maxims:

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

The third category is **relation**, for which Grice offers only one maxim:

1. Be relevant.

Finally, under the category of **manner**, Grice considers *how* something is said, mentioning four maxims:

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief.
4. Be orderly.

Taken together, the maxims suggest that a participant's contribution should be as brief, truthful, relevant, clear, and orderly as the situation permits. If a speaker fails to adhere to any of the general cooperative principles and maxims, the other person is likely to assume that there must be some reasonable explanation and begin a chain of reasoning concerning what might have been meant by what was said. As such, the maxims are like the premises of a logical argument grounding a chain of reasoning. The result of that reasoning process produces a relationship between the maxims and a rationalized understanding of the utterance that Grice identifies as a "conversational implicature."

Conversational implicature may be defined as the most rational guess one can make based on the conversational maxims about what the other person intends by what he or she says in the total context of a conversation. Grice's point is that the speaker assumes that the listener will be able to guess which implicatures are intended based on what is actually said. However, because the sender knows which implicatures are likely to be drawn from what is said, conversational implicature has considerable potential for producing deception. For example, if a student says "I can't get my work turned in on time when my computer is not working," the most rational interpretation is likely to be that the student's computer was malfunctioning and the work was lost or not completed. The student did not say that, of course, but has planted a deceptive trap via the implicature generated in the utterance.

Let us now look at how Grice's theory helps us understand how James Comey might have arrived quite rationally at the conclusion that he was being *directed* to drop the Flynn investigation when President Trump used the *expressive* form rather than the explicit directive verbal form to convey this message. In the context, where the person of higher authority says to a subordinate "I hope you will X," the implicature seems to be that the higher-ranking person is trying to give the order in a polite manner, just as a parent who says, "This room could sure use some dusting" intends to produce the implicature that the child is being instructed to clean the room. Why would one speak implicatively rather than directly? In the Comey example, implicative speaking really seems to be a very efficient way giving an order without the ability to be held responsible for having given that order, especially if, as in this case, the implicative order was not carried out.

ANALYZING THE MACROSEMANTIC/TEXTUAL FOUNDATIONS OF DISCURSIVE DECEPTION

The final level of language where deception may occur is the macrosemantic/textual level (or simply the macrosemantic level). At the macrosemantic level of discourse, we have multi-sentence utterances whose collective meaning can be deceptive. There are four traditional types of macrosemantic discourse that can be used deceptively: descriptions, expositions, narratives, and arguments. This section will briefly explain each type and how it may be used deceptively.

Deception by Disguising Slanted Descriptions as Neutral Ones

Description is the use of language to paint a complex sensory picture of a state of affairs. Its usual purpose is to bring the state of affairs to life in the imagination of the recipient. However, descriptions can be highly variable because one cannot mention every possible thing that might be observable about whatever one is describing. Accordingly, descriptions are always marked by the choices the describer makes about what is important about the thing being described. Thus, the first source of possible deception during description concerns the choices the deceiver makes about what to mention in the description. Moreover, even if two people were to agree on the details of what must be mentioned in their descriptions, they may choose different categorical words to label those features with, for example, one person selecting positively valenced words and the other choosing negatively valenced ones. Thus, descriptions can be intentionally deceptive based on either the details that are chosen for inclusion in the description or on the categorical words chosen to label those details. For example, Hong Kong was described by early explorers as a fragrant harbor (so much so, that its name may be translated literally as "fragrant harbor"). But others who did not want to spill blood fighting over the right to control it in the nineteenth century described

it as "a barren rock," with no maritime value. Both descriptions might properly characterize aspects of Hong Kong over 150 years ago but, depending on one's persuasive perspective, each could be considered to be a deceptive description designed to influence how the British and Chinese governments would treat it in their negotiations with one another.

Deception by Proposing Misleading Expository Taxonomies

Exposition is the process of "exposing" some topic for clearer viewing by breaking it down into its component parts for closer inspection and showing the relationship of the parts to one another. Underlying the expository process is a two-phase activity known as *analysis and classification. Analysis* may be defined as the process of dividing something into its constituent parts *for a particular purpose*. Thus, for example, if someone asks, "How many types of computers are there?" the answer might be "two" (breakdown into parts): "Mac and PC," assuming that the purpose was identifying well-known general operating systems. But a different purpose would offer a different analysis, say: "Laptop and Desktop." And a third analysis with yet a different purpose might say "Portable and Mainframe."

The second process underlying exposition is classification. *Classification* may be defined as the activity of providing category-naming labels to identify the categories that were created during the analysis phase. In the computer example above, the classification process was also illustrated because the labeling process had to be completed before we could even talk about the constituent parts.

All of this might seem to be quite mundane, except that there are very strong possibilities for deception in both the analysis and the classification phases of creating an exposition. For example, suppose someone pointed out in the previous example that there are really three types of general operating systems; Linux had been omitted as if it didn't exist. I could be accused of a deceptive practice in my analysis by ignoring a category I preferred not to mention. If we apply this principle to analyzing one of the rapid social changes occurring in the twenty-first century, suppose someone asks "How many genders are there?" Some will answer "two: male and female." In contrast, *Facebook* offers its users over 50 different gender categories, each with its own classifying name and distinguishing categorical characteristics. So, how many genders are there "really"? It depends on your purpose, which is why the analysis phase of exposition is such a critical factor for understanding the potential for linguistic deception. A deceiver can (mis)represent the complexity of something merely by the number of analytical divisions that are chosen.

The potential for expository deception is equally fraught in the classification phase because each of the analytical divisions must be given a name. As we saw earlier in the chapter, category-naming words are always chosen for a purpose, usually to influence the outcome on some topic. Accordingly, the

exact same principles and examples presented earlier would be relevant here in one of the expository subprocesses.

Deception by Distorting Narrative Elements During Storytelling

Narrative is the discourse form that uses all dimensions of language to tell a story about actors and their actions in a particular set of circumstances. *Narration* may be defined as the process of presenting a sequence of events that involve characters who are doing things in order to accomplish goals, usually in the face of some problems they have to overcome. Accordingly, most fully developed stories have many elements, such as a scene or setting in which a set of characters gets involved in a complicating set of events that require some type of actions by the characters who have various motives for accomplishing their goals; the characters' actions and motives give the story a sense of directionality that leads to a climax of activity, after which there is a resolution or completion of the events that has both consequences for the story participants and a point or lesson for the storyteller and the audience. The potential for deception during storytelling can arise at any and all of these elements of the narrative process. More will be said about deception via narrative in the next section, where its relevance can more easily be seen.

Deception by Offering Misleading Argument Warrants

An argument is a macrosemantic discursive structure composed of two essential parts: a claim and one or more reasons for believing that claim. The relation between a claim and the reason given to support it is called a *warrant*. That is, a warrant is a justification for connecting a claim with the reasons used to support it. Traditionally, two broad types of warrants are recognized: inductive and deductive. Each type comes with its own potential for deception when using arguments to persuade someone to accept a claim as true.

Inductive Warrants

Inductive reasoning is the mental process of turning one's direct or indirect experience with the world into a verbal statement about that experience that one believes to be true. Inductive *warrants*, then, are the principles people use to justify the move from their direct observations to the sentences that they formulate to encode their observations into language. We can identify at least six types of inductive warrants, each with its own potential for use during deception. These are classification, sign, analogy, generalization, causation, and narrative probability warrants.

Classification Reasoning and Warrants. The simplest and yet most fundamentally important type of inductive reasoning is classification. In classification, the person (a) observes an event or object, (b) assigns it to a category

of events or objects, and (c) attaches an appropriate category label to it. What makes classification a significant form of inductive *reasoning* (with a corresponding potential for deception) is that the things we observe do not *announce* which human categories they fit into just because they have been observed. Events occur, things exist, but classifying them is a human activity requiring observers to decide which category the event best fits into, and which word should therefore be used to label it. To "warrant" a move from direct observation to classification, the person must be able to determine the attributes that most significantly characterize the event or object, decide which class or classes those events would possibly match, and then select a label for the classification that they choose to emphasize. Accordingly, deception in the classification process could occur by misidentifying the most significant attributes of the item being classified, by choosing the wrong category for placing the item in, or by self-consciously selecting a misleading label for identifying the item. Thus, this type of inductive warrant makes three additional subtypes of deception possible.

Generalization Reasoning and Warrants. Classification is the essential act of inductive reasoning; however inductive reasoning can go much further than mere classification of a single event. For example, one can use one's specific observations to move to a general summarizing proposition. This is known as *generalization* reasoning because the mind moves from one or more specific observations to a broader claim about more instances that have not been directly observed. Thus, in generalization, the reasoning process claims that what "what is true of a few well-selected instances of something is likely to be true of all other instances as well." Warranting generalizations involves a two-step process. First, one must discover a pattern or characteristic that exists within a number of specific instances of object or event; then based on the pattern, one forms a proposition that predicts that other unobserved instances of the thing will exhibit the same characteristic or pattern. Accordingly, deception in the generalization process could occur if the deceiver chooses a non-representative selection of instances to form the generalization, or if the deceiver makes a prediction that does not take into account additional circumstances that would defeat the generalization that is being claimed.

Sign Reasoning and Warrants. A sign is anything that indicates the existence or presence of something else. The sign is the observed phenomenon; what it is used to predict is the existence of something that cannot itself be directly observed. Thus, the reasoning process runs something like this: The things that can directly be observed to exist are taken as meaning that related things that cannot be directly observed must also exist, have existed in the past, or are coming into existence in the future. Because signs are always component parts of whatever situations they are said to be signs of,

sign reasoning is warranted by identifying the connection between what is taken to be a sign and the situation that is claimed to exist because the sign is actually observed. Accordingly, the deceptive use of sign warrants may involve misleading the other about which signs have actually been observed or how those signs are connected to the situation that the deceiver proposes that they arise from. Further, deception can occur when a pattern of countersigns is self-consciously ignored, or when a particular sign could arise from several different possible situations, but these alternatives are intentionally ignored as possible explanations for the sign that has been observed.

Analogy Reasoning and Warrants. Two things, processes, situations, or events are analogous whenever they have the same essential form—that is, they share a common pattern of structural elements and relations among those elements. Reasoning by analogy induction occurs, then, whenever one identifies a pattern of features in one thing and then proposes that, because something else displays a number of those same features, the second thing will also share additional predicted (but not actually observed) characteristics of the pattern found in the first.

The key to the warrant of this type of induction is the strength of the pattern of features that are common between the two items. Accordingly, deception in the analogizing process can occur in at least two different ways. First, the deceiver may simply claim to find similarities that are not really there. That is, there really is no common pattern even though the deceiver claims to have found one. Second, the deceiver may ignore important differences between the features of the two items that are more important than the similarities that *do* exist. Here, the deception is based on overlooking major breaks in the pattern that make the two things substantially different from one another.

Causation Reasoning and Warrants. To cause something is to be directly responsible for its happening. Accordingly, when we try to explain events in our world we frequently try to identify the observable forces that were available to bring a particular result about. Reasoning by causality induction occurs when we observe that certain forces exist (or did exist in the past) and that they were capable of bringing about some subsequent situation that we are trying to explain. Thus, the reasoning process runs something like this: Although I did not see the moment of causation, factors that I can observe (or did observe) were available to cause the events I want to explain.

What makes this form of thinking into inductive *reasoning* is that the person observes the existence of the forces themselves but does not see the actual moment of causation in which they may have operated. Without actually seeing the moment of causation, the observer is speculating that the forces that were observed could *in principle* have produced the event in question. Causal reasoning, then, moves from one or more antecedent conditions,

which direct observations show to exist (or to have existed prior to the event being explained), to a proposition stating that those factors were in fact responsible for the event. Accordingly, deception during the causal reasoning process can occur in at least three ways: (a) if the deceiver says that certain forces or conditions existed when they in fact did not, (b) if the deceiver says the conditions were of the type that could have caused the event when they actually could not have produced it, or (c) when the deceiver says any conditions that might have prevented these conditions from producing the event were not in the situation when they were in fact part of the actual mix of conditions that existed at the time.

Narrative Reasoning and Probability Warrants. As mentioned earlier, narrative is a form of discourse in which a story is told to make a point. Whenever someone tells a story, they usually try to tell it coherently, so that all of the details included in the telling seem to be necessary to the unfolding of the story. That is, the details must seem to make sense within the framework of the overall story that is being told. For a detail to be narratively probable, then, means that it fits coherently with the general framework of the story being narrated. When a detail does not fit, one must either question the likelihood of the detail or change the nature of the story that is being told. To reason by narrative probability, then, is to claim that certain propositions must be true (or false) simply because they would (or would not) make sense within the framework of a certain narrative story line. Although reasoning by narrative probability is used every day as we give accounts for our behaviors and decisions, it can most easily be seen in legal discourse, where opposing lawyers present competing stories to the jury and try to convince them that their retelling of the events (namely the retelling that most favors their client's position) is the most coherent and comprehensive (Bennett, 1978). To be coherent means that the details make sense when woven together into a unified story line, and to be comprehensive means that no significant details have been omitted in order to achieve whatever degree of coherence and narrative probability the story conveys. Accordingly, deception when using narrative probability can occur if details are omitted that would degrade the receiver's sense that the deceiver has told a narratively probable and coherent tale.

Deductive Warrants
Deductive reasoning is the process of determining what the covert implications are of the sentences we already believe to be true. That is, deductive reasoning moves from the presumed truth of one or more previously formulated beliefs to the creation of a new belief based on *formal relationships that exist among one's old beliefs*. In deduction, no new evidence from the world outside of the language is required in order to reach novel conclusions that apply to the world. The primary requirement is that one manipulates the

information already encoded in a previously formed set of sentences "validly" so that the truth of the original set of sentences guarantees that any new sentences derived through valid deduction must also be true.

The potential for deception during the deductive process comes primarily in two forms. First, the original set of "true" sentences may not all actually be true. That is, the communicator may introduce false information into the deductive set, and therefore even if the sentences are organized and manipulated "validly," the conclusions derived may be false. The second route to deception or error in deductive systems comes when the information expressed in true sentences is manipulated in non-valid ways. In this case, the conclusions arrived at may be true or false, but even if they are true, they are not *guaranteed* to be true by the principles of proper deductive warrants.

Conclusion

This chapter has surveyed the potential for deceptive practices at four different levels of language: lexical, syntactic, speech act, and macrosemantic. It has identified over 20 major different ways that language as a system can be used to lead another person astray when it is advantageous for a deceiver to do so. The goal has been to explain the basic principles of each level in sufficient detail so that a framework for analyzing a communicator's deceptive practices can be discovered and understood. It is hoped that this framework will provide a practical basis for uncovering deceptive practices in interpersonal, group, and public communication situations.

References

Austin, J. L. (1962). *How to do things with words.* Oxford, UK: Oxford University Press.

Bennett, L. R. (1978). Storytelling in criminal trials: A model of social judgment. *Quarterly Journal of Speech, 64,* 1–22.

Brown, R. (1958). How shall a thing be called? *Psychological Review, 65*, 14–21. Reprinted in Brown, R. (Ed.). *Psycholinguistics: Selected papers by Roger Brown* (pp. 3–15). New York, NY: Free Press.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York, NY: Academic Press.

Hayakawa, S. I., & Hayakawa, A. R. (1990). *Language in thought and action* (5th ed.). New York, NY: Harcourt.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago, IL: University of Chicago Press.

Magane, A. (June 21, 2017). Don't tell me Nabra Hassanen, the Muslim girl who was kidnapped outside a mosque and murdered, was a victim of road rage. London, UK. *Independent.* Retrieved November 23, 2017, from http://www.independent.co.uk/voices/nabra-hassanen-islamophobia-ramadan-murdered-hate-crime-road-rage-a7800126.html.

New York Times. (June 8, 2017). Full transcript and video: James Comey's testimony on Capitol Hill. *New York Times*. Retrieved November 23, 2017, from www.nytimes.com/2017/06/08/us/politics/senate-hearing-transcript.html.

Powers, J. H. (2016). *Public speaking: The lively art* (2nd ed.). Singapore: McGraw-Hill Education (Asia).

Searle, J. R. (1979). The classification of speech acts. *Language in Society, 8,* 137–151.

Steinmetz, K. (April 12, 2017). United Airlines 'reaccommodated' a passenger. Is that the euphemism of the year? *Time*. Retrieved November 23, 2017, from time.com/4735588/united-airlines-video-overbooked-reaccommodate-oscar-munoz/.

Deception and Language: The Contextual Organization of Language and Deception (COLD) Framework

David M. Markowitz and Jeffrey T. Hancock

The relationship between deception and language is complex and often inconsistent across studies. For example, prior work has observed that liars tend to use fewer self-references (e.g., *I*, *me*, *my*) than truth-tellers when discussing their views on abortion (Newman, Pennebaker, Berry, & Richards, 2003), but when writing fake hotel reviews, liars use more self-references than truth-tellers (Ott, Choi, Cardie, & Hancock, 2011). Inconsistent patterns of deceptive language have been observed for other cues as well, including negative emotion terms (Burns & Moffitt, 2014; Dzindolet & Pierce, 2004), the number of details in a false statement (Elntib, Wagstaff, & Wheatcroft, 2015), and the frequency of words that suggest complex thinking (e.g., exclusive terms, such as *but* and *rather*; Bond & Lee, 2005; Schober & Glick, 2011). Why is the impact of deception on language mixed?

One possibility is that deception does not affect language patterns and extant results reflect random noise. Given that the number of lies told per day is small (e.g., approximately two; DePaulo, Kirkendol, Kashy, Wyer, & Epstein, 1996; Serota & Levine, 2015), language may not be affected by dishonesty because its signal is not robust compared to truthful discourse.

D. M. Markowitz (✉)
School of Journalism and Communication, University of Oregon, Eugene, OR, USA
e-mail: dmark@uoregon.edu

J. T. Hancock
Department of Communication, Stanford University, Stanford, CA, USA
e-mail: jeff.hancock@stanford.edu

On the other hand, deception may affect language, but not uniformly. A recent meta-analysis provides evidence for this claim. Hauch, Blandón-Gitlin, Masip, and Sporer (2015) collected 44 studies that used automated methods to analyze deceptive language patterns, and the data revealed small to moderate effect sizes for cues that betray deceit. Importantly, however, these effects were influenced by moderators, such as the interaction type (e.g., if no interaction occurred, if there was an interview) and the production mode (e.g., written, spoken, or typed communication). The moderators, which often changed the direction and magnitude of the effect sizes for many cues in the meta-analysis, suggest that aspects of a situation matter when investigating the relationship between deception and language.

In this chapter, we draw on empirical moderator analyses from Hauch et al. (2015) and theories from deception research (Buller & Burgoon, 1996; Levine, 2014) to argue that *context* is an important aspect of each deception and requires consideration when making predictions about how lying affects language (and to understand anomalous results). Context is a vague term, however, and has received limited treatment from the deception literature. We dissect what context means for deception, the role of context in language production independent of deception, and how context plays a role in the relationship between deception and language. This approach integrates research from psychology, communication, and linguistics, concluding with a framework to understand how context affects deceptive word patterns, called the Contextual Organization of Language and Deception (COLD) framework.

Does Deception Affect Language?

The idea that deception affects language is supported by evidence suggesting that word patterns are modified by social and psychological processes. For example, a long tradition of clinical research has observed that depressive individuals are often more emotional than non-depressive individuals (Beck, 1967). The writing styles of people who rate highly on depression scales (Rude, Gortner, & Pennebaker, 2004) and who die of suicidal causes (Markowitz & Hancock, 2017) reveal a consistent pattern, as markers of emotion (e.g., positive and negative affect; *amazing* and *horrible*, respectively) are often amplified in the language patterns of people who are experiencing a distressing psychological event.

A second example highlights how academic success in college can be forecasted by students' admissions essays before they were accepted to the university. Pennebaker, Chung, Frazee, Lavergne, and Beaver (2014) analyzed the text of over 50,000 admissions essays and observed that better academic performance in college was associated with a more analytic writing style. Students with higher GPAs at the end of four years wrote with more articles (e.g., *a*, *the*) and prepositions (e.g., *above*, *below*), which reflect critical thinking, than students with lower GPAs. The lower-achieving students used

more pronouns (e.g., *I*, *me*) and verbs (e.g., *can*, *has*) than higher-achieving students, which reflects a narrative thinking style that is less analytic and complex. Together, the prior examples suggest that language reflects important information about who we are, what we are thinking and feeling, and what we are experiencing psychologically (Pennebaker, 2011). It is reasonable to assume, then, that language can be affected by deception and reveal distinct verbal patterns compared to truthful statements.

How does deception affect word patterns? Recall, prior meta-analytic work by Hauch et al. (2015) revealed small to moderate effect sizes for the effect of deception on language variables when looking across studies. For example, lies contained fewer words and more negative emotions compared to truths, but most patterns were influenced by moderators. The Hauch et al. (2015) meta-analysis accounted for five primary moderators that change the relationship between deception and language: (1) *event type* (e.g., a first-person experience, reporting of attitudes), (2) *valence of the deception* (e.g., if the false topic is positive, negative, or neutral), (3) *interaction type* (e.g., if no interaction occurred, the interaction took place online, there was an interview, or a face-to-face interaction occurred), (4) *motivation to lie* (e.g., no motivation, low to medium motivation, high motivation), and (5) *production mode* (e.g., handwritten, typed, spoken text).

The moderator analyses help to understand why language effects may be inconsistent for different deceptions. For example, across studies, lies often contain fewer words than truths, but considering the interaction type as a moderator for the effect reveals a different relationship between deception and language. In a one-way interview setting, a dyadic interpersonal setting, or if no interaction occurs, liars typically communicate with fewer words than truth-tellers. If the interaction occurs online (e.g., instant message chat, email), however, liars use *more* words than truth-tellers. These data suggest that simple changes to the communication format can modify deceptive language patterns and accounting for contextual elements is fundamental to achieve a more complete understanding of how deception changes behavior.

A second example considers how deception affects the rate of first-person singular pronouns (Newman et al., 2003; Pennebaker, 2011). Across studies, Hauch et al. (2015) reported a null effect of deception on first-person singular pronouns, which is unsurprising given that study-by-study comparisons find inconsistent results of lies containing both amplified (e.g., Ott et al., 2011) and attenuated rates of self-references (e.g., Larcker & Zakolyukina, 2012; Newman et al., 2003) relative to truths. Considering the valence of the deception as a moderator for the effect, however, clarifies the relationship. When the situation is negative (e.g., discussing a crime), lies often contain fewer first-person singular pronouns relative to truths, possibly as a psychological distancing mechanism (Newman et al., 2003). On the other hand, when people talk about a neutral event (e.g., writing about a hotel stay that never occurred; Ott et al., 2011), lies contain more first-person singular pronouns than truths, possibly to increase the speaker's credibility.

The evidence by Hauch et al. (2015) suggests that deception affects language, but the relationship is complex given the important moderator effects. These findings proffer that a universal approach to understanding how deception affects language is clearly problematic. Yet, to date, context has not been systematically modeled for the relationship between deception and language, at least in the literature on computerized text analysis. Nonetheless, an examination of the papers from Hauch et al. (2015) revealed that nearly half of the studies cite contextual factors or "the context" when anomalous results are observed compared to prior research or expectations. What does context mean for the relationship between deception and language? In the following sections, we identify how deception and language are both context-contingent phenomena. Finally, we propose three features of context—psychological dynamics, pragmatic goals, and genre conventions—that influence how deception affects language.

Context for Deception

Deception theories sometimes consider contextual factors that affect deception detection. Two frameworks, Truth-Default Theory (Levine, 2014) and Interpersonal Deception Theory (Buller & Burgoon, 1996), both articulate how context can play a role in deception detection studies.

Truth-Default Theory

Levine's (2014) Truth-Default Theory (TDT) proposes that people are poor lie detectors because they assume that others are predominantly honest, independent of message veracity. TDT argues that deception detection efforts can improve, however, when people listen to what is said (e.g., the communication content) and absorb information in context. Levine (2014) suggests that context refers to the specific communication act and other situations relevant to the communication act. For example, if two American college friends meet in the dining hall for lunch at their university to discuss spring break plans, the current situation refers to the collocated discussion about the spring break trip. Relevant situations for the communication content are prior conversations that may have occurred (e.g., sorting through options of where to go) and new information that one person holds (but has yet to tell the other person) about the trip. Considering *content in context*, as Blair, Levine, and Shaw (2010) suggest, can help to improve lie detection accuracy and expose parts of a deception that may lead to improved veracity judgments.

What does content in context provide to the lie detector? First, content in context helps to highlight contradictions that the liar may communicate. For example, friends often try to manage the impressions of others by lying about their availability over text messaging (Hancock, Birnholtz, Bazarova, Guillory, Amos, & Perlin, 2009). If a message sender was unresponsive

because of poor cell service but he or she was also active on social media, taking content in context (e.g., inconsistencies between offline and online behavior) can help to uncover their dishonesty. Indeed, research suggests that contextual features are often used in the actual detection of everyday lies (Levine, Park, & McCornack, 1999), including information from third parties that is inconsistent with the deceptive statements (e.g., knowing from a Facebook post that a friend did not stay home, but went to a party instead).

Second, content in context provides cues to suggest what is normal or possible in a situation. That is, content in context helps to differentiate false exaggerations (e.g., "I just ate a million donuts") from harmful lies and allows people to focus on messages that are suspicious. Finally, content in context can provide idiosyncratic information that suggests if deception is plausible for a given situation. This element of content in context is especially applicable to niche areas of deception, including politics or finance, because people who have expertise or inside knowledge about a subject area can detect deception better than outsiders (Blair et al., 2010; Levine, 2014).

The value of using content in context for deception detection was demonstrated in several studies by Blair et al. (2010), who had participants watch videotaped interrogations with contextual information (e.g., a case file on the suspect) or without contextual information. The data revealed that both professionals and students were better lie detectors when presented with contextual information than without contextual information. Their veracity judgments exceeded chance, and there were fewer "false alarms" when contextual information was given relative to when it was not provided. The work by Blair et al. (2010) and Levine (2014) suggests that content in context can aid in deception detection because the situation provides clues about contradictory, possible, or specialized data related to a possible act of deception. To best understand a deception, contextual factors should be evaluated.

Interpersonal Deception Theory

Interpersonal Deception Theory (IDT; Buller & Burgoon, 1996) proposes that relationship dynamics in a deception are context dependent and contextual features of a situation modify deception detection abilities. These include but are not limited to: (1) the degree of interactivity between two communicators, (2) attributes of each communicator that affect how the interpersonal interaction will unfold (e.g., personality, goals), (3) information and behavioral familiarity, (4) the affective relationship of the communicators before the interaction occurs, and (5) honesty expectations at the onset of the interaction. At its core, the theory treats deception as a negotiation between the sender and receiver; therefore, understanding interpersonal relationship dynamics (e.g., roles of each communicator, emotional valence of the situation) can help to diagnose why lies are undetected.

Experimental evidence from Burgoon and colleagues supports the idea that deception detection is affected by contextual features of a situation,

especially parts of the interpersonal relationship. Burgoon, Buller, and Floyd (2001) manipulated the level of interactivity between two communicators in an interpersonal setting. Participants experienced either high interactivity (e.g., engaging in a dialogue) or low interactivity (e.g., a monologue, where one person communicates to another without message exchange), and lied or told the truth to their partner when discussing four topics: (1) "Tell about the most significant person in your life," (2) "Tell about a mistake you made recently," (3) "Describe the most unpleasant job you have ever had to do," and (4) "Talk about responsibility." The data revealed that message receivers were less accurate detecting deception under high interactivity settings relative to low interactivity settings. Burgoon et al. (2001) suggest that receivers may be less willing to suspect or detect deception interpersonally because people are incorrectly biased by cues (e.g., mutuality), but message senders are also more likely to control how information is communicated. Therefore, context constrains how people communicate a deception and their ability to detect false statements.

Other studies in the IDT tradition suggest that interpersonal deceptions are context dependent because communicators influence conversations to make lies appear similar to truths. If a liar suspects that the message receiver is catching on to the deception, he or she may adjust future communication patterns to avoid detection (see Vrij, 2008). It is crucial to understand how contextual elements affect deception because most lies are interactive and involve two people. Each deception, therefore, demands a unique communication style and is associated with relationship dynamics that constrain how the lie will be communicated.

Together, Truth-Default Theory (Levine, 2014) and Interpersonal Deception Theory (Buller & Burgoon, 1996) suggest that deception detection is difficult because we believe that most people are honest, people often do not use content and context-related cues to detect lies, and dynamics of interpersonal relationships can influence how a lie is told and perceived. These theories suggest that context matters for lie detection, but they also guide how context matters for deceptive message production as well. To investigate how context affects the relationship between deception and language, the next section identifies how language patterns are also context contingent, independent of deception.

CONTEXT FOR LANGUAGE

According to Applegate and Delia (1980), communication messages are affected by five features of context: the physical setting (e.g., the location of the communication act), the social and relational setting (e.g., the relationship dynamics of the communicators), the institutional setting (e.g., an establishment with a specific purpose, such as an office, school, home), the functional setting (e.g., the reason for communication), and the cultural setting (e.g., nationality). The crossing of communication settings is called a

situation (Burleson, 2009). The prior example of students planning a spring break trip can be categorized into Applegate and Delia's (1980) settings: Two American (e.g., cultural setting) college friends (e.g., social and relational setting) meet in the dining hall (e.g., physical setting) for lunch at their university (e.g., institutional setting) to discuss plans for spring break (e.g., functional setting).

Applegate and Delia (1980) provide a useful foundation for understanding context because their model suggests that interpersonal and environmental characteristics contribute to how a communication act occurs. Words are not produced in a vacuum, but they are dependent on and affected by social, institutional, and cultural influences to meet the expectations of other communicators (Levelt, 1989). People write emails to colleagues differently than they write emails to friends, send text messages to a boss differently than a significant other, and write newspaper articles differently than love letters. Communication works because people recognize setting cues and adjust their behavior when situations change (Clark, 1996; Levelt, 1989).

The Applegate and Delia (1980) framework can be supplemented, however, by considering the communication genre as a setting that affects language production. Knapp, Daly, Albada, and Miller (2002) suggest that message-related variables (e.g., the linguistic style, the source, and the audience) are important contextual elements for any communication act because communities have norms that shape how people produce word patterns (Biber, Connor, & Upton, 2007). For example, Tweets are constrained to a character limit and blogs often do not contain a word maximum. Each medium has different conventions that shape a discourse community. Consequently, we argue that an important contextual element of communication is the genre, which captures how people talk according to community norms (Biber et al., 2007). The genre typically changes when deceptions change, suggesting that this contextual element is important for understanding how deception may affect language patterns.

Together, the physical setting, social and relational setting, institutional setting, functional setting, cultural setting, and the genre form a situation and influence how language is communicated. This idea is most clearly demonstrated in interpersonal communication, as Burleson (2009) suggests that certain features of context "shape and may even mandate" that people communicate according to dynamics of a community, a social relationship, or their own goals (p. 157). When people abide by elements of genre or their discourse community, they reinforce the social structures that allow communication to occur (e.g., speech acts, turn-taking, nonverbal gestures; Burleson, 2009). Therefore, language has structure that is "beyond the sentence" (Biber et al., 2007, p. 8) and recognizing this structure or context is reflexive. People do not need to think about how to speak to their boss or someone of high social status because contextual cues about social relationships are built into the fabric of a situation (e.g., previous interactions, social expectations that suggest how rank affects communication style; Burleson, 2009).

These findings from linguistics and psychology make clear how context plays a fundamental role in how language is used. In the next section, we integrate these observations with research on deception to develop a context-contingent framework for the relationship between deception and language.

The Contextual Organization of Language and Deception (COLD) Framework

What contextual elements of a deception matter when evaluating how deception affects language? Prior research has often considered emotional and cognitive variables to understand the psychological dynamics of a situation inferred through language. For instance, researchers have considered how people express emotional content in Yelp reviews (Margolin & Markowitz, 2018), how emotion can be transferred across social networks (Kramer, Guillory & Hancock, 2014), and how people feel after experiencing a trauma (Cohn, Mehl, & Pennebaker, 2004). Empirical work also suggests that cognitive information or data related to cognitive load and complexity (see Sporer, 2016 and Vrij, 2008 for a review) also reveal how people may construe a situation. With substantial evidence suggesting that psychological dynamics matter when people communicate, it is reasonable to suggest that these elements will also matter for deception.

Communication goals, or the reasons why someone participates in a communication act, are also contextual and vary across deceptions. For example, evidence suggests that goals change how people choose to communicate in an online setting. Bazarova and Choi (2014) examined how people use different social media formats (e.g., Facebook status updates, wall posts, or private messages) depending on their social and relationship goals. For intimate conversations, people use private messages rather than status updates or wall posts. These data propose that people approach a situation with objectives, and communication behavior (e.g., language) is often a reflection of these goals. We suggest that deception goals, which vary by situation (Buller & Burgoon, 1996; Turner, Edgley, & Olmstead, 1975; Vrij, 2008), will modify language patterns because the reasons for communicating a deception vary widely.

Finally, it is important to consider how communication behavior reflects the environment that it is situated within. The genre is an important contextual variable for the relationship between deception and language because discourse patterns often reflect behavior that is typical for a given setting (Burleson, 2009). A deception is situated within an environment, which has its own conventions (e.g., text messages are communicated differently than political speeches), and is constrained to language patterns that reflect the genre's norms. For deception, the genre is important because people who speak outside of the norms of the discourse community may be viewed as suspicious (Levine, 2014). People want to avoid detection and speaking in a genre-consistent manner is essential to blend in.

Based on these observations, we propose a model for assessing the contextual effects of deception on language: the Contextual Organization of Language and Deception (COLD) framework. The COLD framework suggests that there are three important aspects of context for any deception: (1) the psychological dynamics, or the emotional and cognitive elements of the lie, (2) the pragmatic goals of the speaker, or what he or she is trying to accomplish with the deception, and (3) the communication conventions in which the deception takes place, including the genre of the discourse community. We discuss these items as contextual constraints that should be considered when making assessments of how deception will affect language.

Psychological Dynamics

Psychological dynamics consists of two key dimensions. The first dimension reflects how the liar's emotional experience is different from a truth-teller's emotional experience. Liars often try to approximate the genuine emotions that truth-tellers feel, and there is mixed evidence for the effect of deception on emotion cues (e.g., negative emotion terms). For example, Burns and Moffitt (2014) evaluated false and truthful calls to a 9-1-1 dispatcher and observed that lies contained fewer negative emotion terms than truths. The authors suggest that truth-tellers felt genuine fear compared to deceptive callers who could not approximate similar levels of distress. On the other hand, liars can also experience more negative emotions than truth-tellers for reasons including worry of being detected, guilt for deceiving another, or breaking a moral code (see Ekman, 2001; Vrij, 2008). As offered by Hauch et al. (2015), emotional language use affected by deception is moderated by the valence of the topic, the type of interaction, motivation, and the production mode. Therefore, considering the impact of deception on emotions requires contextual rather than universal treatment.

A second dimension of psychological dynamics considers the cognitive experience of the deceiver. Some research suggests that false narratives contain fewer cognitive complexity markers than truthful narratives (e.g., exclusive terms; Schelleman-Offermans & Merckelbach, 2010), while other studies find that cognitive complexity markers are unrelated to deception (e.g., in the text of an online dating profile; Toma & Hancock, 2012). Considering context and the Hauch et al. (2015) moderators highlights that the interaction type of a deception may impact the frequency of exclusive terms. For example, exclusives are produced significantly less often in lies relative to truths for interviews, face-to-face interactions, and situations without an interaction. The effect disappears, however, in online environments. Considering features of the situation, particularly the interaction type from the prior example, can therefore help to understand why deception and language results may be inconsistent with empirical findings.

Together, emotion and cognition variables are important contextual elements of a deception because they reflect how the deceiver responds

psychologically to an act of lying. Prior work has observed that psychological variables change when people respond to events that deviate from everyday experiences (e.g., distress, death; Pennebaker, 2011). We argue that deception is not unique in this respect and each deception should be treated as a situation that creates specific psychological dynamics (e.g., emotional and cognitive demands) for the communicator.

Pragmatic Goals

Communication goals can broadly be separated into two types: primary and secondary goals (Dillard, Segrin, & Harden, 1989). Primary goals often represent the reasons for influencing another person or their behavior (e.g., to produce a false belief in another), and secondary goals are motivations that support the primary goal (e.g., to prevent embarrassment). Turner et al. (1975) provide a specific set of motivations related to deception, including: to save face, guide the social interaction, avoid tension or conflict, influence or control the situation, and increase interpersonal power over another. Given that there are different strategies and reasons for lying, another important feature of context considers what the liar is trying to accomplish and how this goal is reflected in language.

A relevant example considers how communication goals modify deceptive language patterns. In 2012, former social psychologist Diederik Stapel was convicted of data fraud in over fifty research publications. Markowitz and Hancock (2014) analyzed Stapel's first-authored fraudulent and genuine papers, finding that he overused science-related terms in his fake relative to genuine papers. Presumably, Stapel attempted to make the deceptive reports appear credible or as credible as the genuine reports, but he inappropriately estimated the frequency of words related to means, methods, and investigation. Considering the reasons for a deception can help to understand why language patterns are different between false and truthful corpora.

The context argument here suggests that people have different reasons for lying and use communication patterns to match their deception. The effect of deception on language should not be uniform because liars have different goals. Crucially, categorizing lies by deception goal may help to understand why language patterns are consistent or inconsistent across studies.

Genre Conventions

A genre, or the discourse community of a speaker (Biber et al., 2007), influences the relationship between deception and language by first constraining a person's language to a refined set of features for conversation. Then, these genre-normative language features are modified by deception. This approach, where a person's discourse is first constrained by communication settings (Applegate & Delia, 1980; Knapp et al., 2002) and then false intentions alter language patterns, is fundamental to the context argument.

Take again the example of scientific fraud, where the genre (science communication) has written, edited, and impersonal writing style conventions. Language dimensions such as first-person singular pronouns (e.g., *I, me*), which are scarce in science writing, are unlikely to be affected by deception because they are generally unconventional in the genre. The characteristics of science writing, as suggested by a COLD framework perspective, are different from other genres that often include personal pronouns such as politics (Pfiffner, 1999). Therefore, first-person singular may be an important language feature in one genre (e.g., political speeches) but not in another (e.g., science writing) because the community conventions, independent of deception, are unique. The COLD framework argues that the effect of deception on language will be influenced by the genre of the deception, and the robustness of the effect will likely be limited to that genre.

Taken together, the COLD framework suggests that there are three contextual factors that need to be considered when evaluating the relationship between deception and language. A universal conceptualization of how deception affects language will likely miss how each deception's language effect is determined by the psychological dynamics of the deception, the goals of the communicator, and the genre in which it is communicated. By considering the psychological dynamics (e.g., emotional and cognitive processes affecting language production), pragmatic goals (e.g., what the liar is attempting to accomplish and how this is reflected linguistically), and genre conventions (e.g., features of the genre or discourse community that constrain language use), a more systematic model of how deception affects language should arise and allow for predictions to be made across studies. Below, we apply the COLD framework to a database with systematic differences in context by examining US presidential deceptions.

Political Deception

Does deception affect political speech? We approach this question by applying the COLD framework to six deceptions from US Presidents: George W. Bush and the War in Iraq, Lyndon B. Johnson and the Gulf of Tonkin Incident, Bill Clinton and the Monica Lewinsky Affair, Richard Nixon and Watergate, John F. Kennedy and the Cuban Missile Crisis, and Ronald Reagan and the Iran Contra Affair (see Alterman, 2004).

To apply the COLD framework to these data, we first hold genre constant by comparing all speeches within political discourse. Second, we organize presidential lies by the pragmatic goals each president was trying to accomplish with their deception. Research in political science (Alterman, 2004; Pfiffner, 1999) has observed that political deceptions can be broadly arranged into policy lies (e.g., reasons for bringing a country to war; George W. Bush, Lyndon B. Johnson), lies to prevent embarrassment (e.g., marriage infidelity or illegal campaign tampering; Bill Clinton, Richard Nixon), and state secrets (e.g., lying by omission or concealment to prevent domestic or international

chaos; John F. Kennedy, Ronald Reagan). We expect that these pragmatic goals will modify language between goal types (e.g., policy lies should reveal a different language style than lies to prevent embarrassment).

Third, we consider how deceptive language patterns are affected by the psychological aspects of a deception. According to the Newman Pennebaker (NP) Model of Deception (Newman et al., 2003), which was not developed for political speech, but its features have been substantiated across a number of empirical deception studies (Bond & Lee, 2005; Hauch et al., 2015; Toma & Hancock, 2012), liars betray several deception cues with language. For example, liars tend to use fewer self-references (e.g., *I*, *me*) than truth-tellers to focus the attention away from the self and onto other objects in the situation. Liars also use reduced rates of exclusive terms (e.g., *but*, *unless*) than truth-tellers, as a reflection of the reduced cognitive complexity often associated with telling a false story. Lies subsequently have less detailed information than truths because it is difficult for people to tell a detailed story that is fabricated (Markowitz & Hancock, 2014). Further, lies typically contain more negative emotion terms (e.g., *hate*, *dislike*) than truths to reflect the distress and anxiety associated with telling a false statement (Ekman, 2001). Finally, motion terms (e.g., *change*, *follow*) are often more prominent in lies relative to truths as they can help to move a false story forward and distract the listener from detecting deception.

Here, we perform an exploratory analysis to investigate how lies from the genre of politics, arranged by deception goals, modify rates of first-person singular pronouns, exclusive terms, negative emotion terms, and motion terms (Newman et al., 2003).

METHOD

The Center for Public Integrity (CPI, 2008), a nonprofit group of investigative journalists, created a database of public statements by eight Bush administration officials about the rationale for the Iraq War (President George W. Bush, Vice President Dick Cheney, Secretary of State Colin Powell, National Security Adviser Condoleezza Rice, Defense Secretary Donald Rumsfeld, Deputy Secretary Paul Wolfowitz, White House Press Secretaries Ari Fleischer and Scott McClellan). CPI researchers gathered transcripts from September 11, 2001, to September 11, 2003, and identified objectively false claims. Together, 531 texts were collected and divided into false (49,797 words) and truthful control statements (139,200 words). Some false statements ($n=24$) did not contain a truthful control, resulting in a database of 1,036 statements.

We acquired transcripts for the other five presidents from the American Presidency Project (Woolley & Peters, 2009), an archive containing over one hundred thousand presidential documents. Consistent with the CPI's methodology to investigate the Bush administration's utterances, we analyzed a president's transcripts from the timeframe during which event-specific

lies were told. A total of 111 statements were gathered across the five presidencies, and statements were divided into their deceptive (Johnson $n=22$; Clinton $n=8$; Nixon $n=12$; Kennedy $n=3$; Reagan $n=10$; 5750 total words) and truthful statement types (Johnson $n=23$; Clinton $n=8$; Nixon $n=12$; Kennedy $n=3$; Reagan $n=10$; 142,680 total words).

Text Processing

We used Linguistic Inquiry and Word Count (LIWC; Pennebaker, Booth, & Francis, 2007) to assess false and truthful control statements by speaker. LIWC is a well-validated tool that calculates the frequency of a single word across its internal dictionary of social dynamics, psychological processes, and parts of speech (see Hauch et al., 2015). For example, the statement from Lyndon B. Johnson, "I was later informed that the ships or the unidentified vessels continued to approach our two destroyers and they opened fire," contains 21 words and LIWC increments each category as a percentage of the total word count. Since there was one instance from each of the NP model categories, first-person singular pronouns (e.g., *I*), exclusives (e.g., *or*), negative emotion terms (e.g., *destroyers*), and motion terms (e.g., *approach*) represent 4.76% (1/21) of the input text. NP model dimensions were drawn from the standard LIWC 2007 dictionary and compared across false and truthful statements for all presidents and officials.

Results

The data were analyzed using hierarchical linear mixed models with statement type (false versus control) as a between-subjects factor. We controlled for data non-independence by nesting statements within speaker and entering this variable as a random effect into each linear mixed model.

If deception affects language in a context-contingent manner, as predicted by the COLD framework, we would expect significant interaction effects for statement type (deceptive vs. truthful control statements) and pragmatic goals (policy lies, lies to prevent embarrassment, state secrets). Consistent with the COLD framework and the idea that deceptive language patterns are context dependent, there were significant interaction effects between the pragmatic goal and statement type for several NP dimensions: first-person singular pronouns [$F(2, 1128.36)=28.38$, $p<.001$], exclusives [$F(2, 1128.99)=3.27$, $p=.038$], and negative emotion terms [$F(2, 1133.09)=11.00$, $p<.001$], but not for motion terms [$F(2, 1131.03)<1$].

As expected by the contextual approach of the COLD framework, the language effects were consistent within pragmatic goals and different across pragmatic goals (see Fig. 10.1). The combined results for policy deceptions (the Bush administration and President Lyndon B. Johnson; the left panel of Fig. 10.1) revealed patterns consistent with typical NP model predictions

for first-person singular pronouns [$F(1, 1128.41) = 63.63$, $p < .001$], exclusive terms [$F(1, 1129.18) = 11.54$, $p = .001$], and negative emotion terms [$F(1, 1133.16) = 423.51$, $p < .001$]. Motion terms were significantly different across statement type, but in the opposite direction as predicted by the NP model [$F(1, 1131.27) = 37.98$, $p < .001$]. Together, the data suggest that statements from the Bush administration and President Lyndon B. Johnson had speaking patterns largely consistent with the NP model (e.g., fewer self-references and exclusive terms, more negative emotion terms) in false compared to truthful control statements, when holding the speaker constant (see Fig. 10.1 for univariate effects by speaker).

The combined results for lies to prevent embarrassment (Presidents Clinton and Nixon; the middle panel of Fig. 10.1) were less clear but markedly different from the policy deception language effects. That is, lies contained *more* first-person singular pronouns [$F(1, 1128.36) = 37.49$, $p < .001$] and a trend toward more exclusive terms [$F(1, 1128.98) = 2.59$, $p = .108$] compared to truths. Negative emotion and motion terms were in the opposite direction of the NP model and also failed to reach significance as well [Fs < 1]. Finally, combined results for Presidents Kennedy and Reagan (the right panel of Fig. 10.1) revealed no significant differences across NP model features [Fs < 1.1]. Together, the mixed language patterns between pragmatic goals but consistent language patterns within pragmatic goals suggest that this element of context plays a crucial and systematic role in the way that a deception is communicated linguistically.

This study has several strong effects that differ from typical NP model predictions. For example, while the rate of first-person singular pronouns was lower in lies relative to truths for policy deceptions (Presidents Bush and Johnson), more self-references were used in the lies relative to truths for lies to prevent embarrassment (Presidents Clinton and Nixon). How can we reconcile such mixed effects? The COLD framework offers a foundation to evaluate how deception affects language across two crucial dimensions, with genre held constant: (1) psychological dynamics, and (2) pragmatic goals. Recall, the psychological dynamics of a lie involve how a person responds emotionally and cognitively to the deception. Considering the psychological dynamics of embarrassment lies that are personal in nature (e.g., dishonesty about marriage infidelity), they often affect a person's face, or his/her positive sense of self (Goffman, 1959). An increase in first-person singular pronouns may be a psychological process used to augment speakers' authenticity or credibility because their reputation has suffered (Ott et al., 2011). This can be accomplished by using more self-references and more concrete language (see Larrimore, Jiang, Larrimore, Markowitz, & Gorski, 2011; Pennebaker, 2011). On the other hand, lies that are less instrumental for the self and mainly concern others (e.g., policy deceptions; Pfiffner, 1999) may focus more on the collective and less on the speaker to diffuse responsibility and divert attention from the self. Our data are consistent with this possibility,
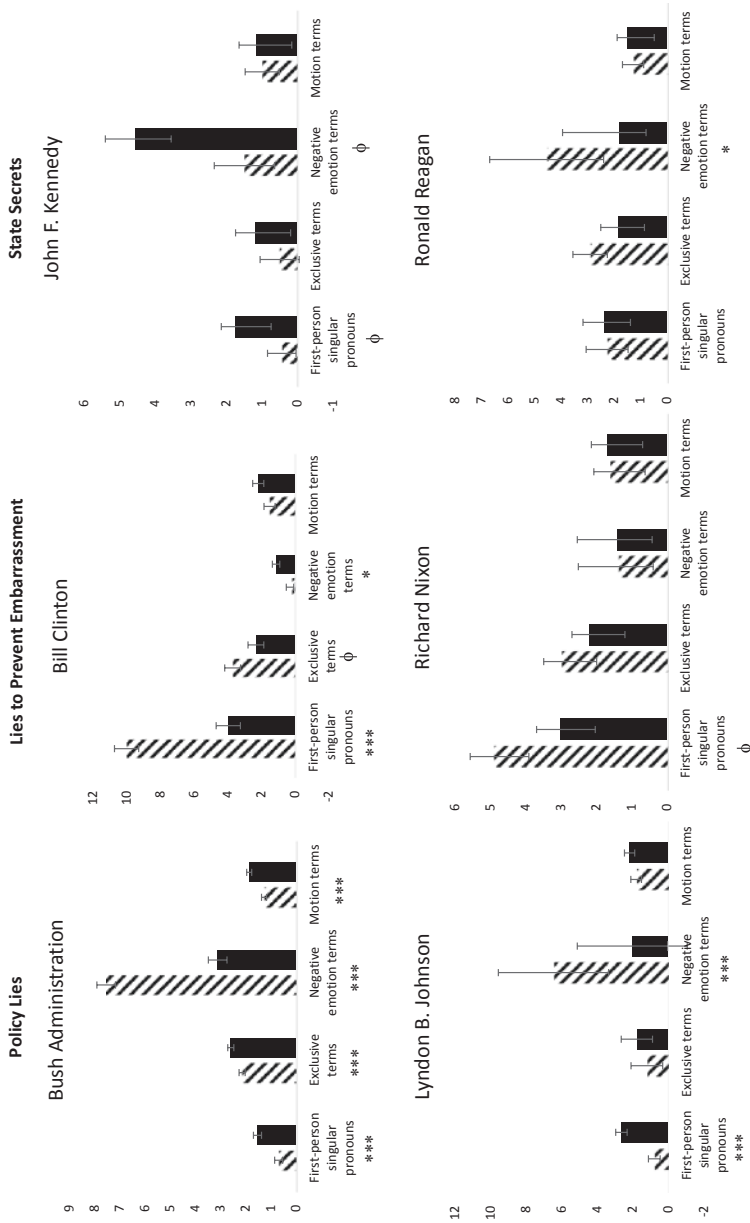
Fig. 10.1 Vertical axes represent raw LIWC percentages of the total word count. Striped bars indicate deceptive statements and solid bars indicate truthful control statements. * $p < .05$, *** $p < .001$, $\phi = p < .08$. Error bars are Standard Errors

with President Lyndon B. Johnson using a lower rate of first-person singular in lies relative to truths, suggesting that this distancing was likely a psychological response to manage the lie.

The pragmatic goals associated with policy lies and lies to prevent embarrassment are also not uniform, and therefore, we should not expect the language patterns reflecting these deception goals to be consistent across different lies. That is, the goal to convince a country of a war effort is qualitatively different than saving face from a marriage scandal. These unique goals should shape the manner in which deception affects language and we found evidence of this contention, similar to how a fraudulent scientist (Markowitz & Hancock, 2014, 2016) should deceive differently than someone writing a fake online dating profile (Toma & Hancock, 2012) because the purpose for the deception is distinct. A fraudulent scientist may fake data and write his or her report differently than a genuine scientist to achieve recognition in academia, while an online dater may lie to portray a more idealized self (see Markowitz & Hancock, 2018). With unique deception goals across settings, the effect of deception on language should not be universal and should be considered a context-contingent phenomenon.

Together, these data and the COLD framework provide empirical and theoretical evidence that political deceptions are not told with uniform language patterns. Psychological dynamics and pragmatic goals of each deception (e.g., policy lies, lies to prevent embarrassment, state secrets) systematically influenced how deception affected language.

Conclusion

This chapter argues that the effect of deception on language is context dependent, and after considering what context means for deception and language, three features of a contextual framework were developed. First, psychological dynamics considers how deception affects the emotional and cognitive experience of the communicator, as reflected in language (Hauch et al., 2015). Second, pragmatic goals suggest what the liar is trying to accomplish with the deception. Because goals shift based on setting features (Applegate & Delia, 1980; Burleson, 2009; Knapp et al., 2002), deception should not affect language uniformly across lying situations. Finally, genre conventions suggest how the influence of deception on language is first constrained by the genre and the discourse community of each situation (Biber et al., 2007). Considering these three elements together can provide a model of how deception affects language and how communicators change their verbal patterns based on deception-specific characteristics.

We also applied the COLD framework to an untested dataset of political lies and used NP model features to investigate how deceptions modify language patterns (Newman et al., 2003). Separating the deceptions by pragmatic goals provided a clearer depiction of how false political statements

compare to truthful political statements. By positioning deceptive language patterns in the COLD framework, inconsistent outcomes of presidential deceptive speech relative to established theory or prior empirical work can be better understood. Our overarching conclusion suggests that with unique psychological dynamics affecting how a lie is told and distinct motivations for lying within a single genre, the effect of deception on language should not be uniform across studies. We can strategize and learn about the relationship between deception and language by considering the contextual elements that affect how a lie is communicated.

We recognize, however, that our framework does not provide an exhaustive list of dimensions that may matter for the relationship between deception and language. We provide only a starting point for researchers to build off of our theorizing and understand how deception plays an important role in false language production when situations change. Our goal with this chapter was to argue that context deserves additional treatment, positioning, and defining in the deception literature. We offer why deception and language are context-contingent phenomena and explain three characteristics of context that influence how deception affects language. Future research should expand on this model and approach deception as a phenomenon that is modified by psychological dynamics, pragmatic goals, and genre conventions that are unique to each deception.

References

Alterman, E. (2004). *When presidents lie*. New York: Viking Penguin.

Applegate, J. L., & Delia, J. G. (1980). Person-centered speech, psychological development, and the contexts of language usage. In R. S. Clair & H. Giles (Eds.), *The social and psychological contexts of language usage* (pp. 245–282). Hillsdale, NJ: Lawrence Erlbaum.

Bazarova, N. N., & Choi, Y. H. (2014). Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *Journal of Communication, 64,* 635–657. https://doi.org/10.1111/jcom.12106.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects.* New York, NY: Harper & Row.

Biber, D., Connor, U., & Upton, T. A. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure.* Philadelphia: John Benjamins Publishing Company.

Blair, J. P., Levine, T. R., & Shaw, A. S. (2010). Content in context improves deception detection accuracy. *Human Communication Research, 36,* 423–442. https://doi.org/10.1111/j.1468-2958.2010.01382.x.

Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology, 19,* 313–329. https://doi.org/10.1002/acp.1087.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x.

Burgoon, J. K., Buller, D. B., & Floyd, K. (2001). Does participation affect deception success? *Human Communication Research, 27,* 503–534. https://doi.org/10.1111/j.1468-2958.2001.tb00791.x.

Burleson, B. R. (2009). The nature of interpersonal communication: A message-centered approach. In C. R. Berger, M. E. Roloff, & D. R. Roskos-Ewoldsen (Eds.), *The handbook of communication science* (pp. 145–165). Thousand Oaks, CA: Sage.

Burns, M., & Moffitt, K. (2014). Automated deception detection of 911 call transcripts. *Security Informatics, 3,* 1–9. https://doi.org/10.1186/s13388-014-0008-2.

Center for Public Integrity. (2008). *'The Iraq War Card' project.*

Clark, H. H. (1996). *Using language.* Cambridge, UK: Cambridge University Press.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science, 15,* 687–693. https://doi.org/10.2307/40064028.

DePaulo, B. M., Kirkendol, S., Kashy, D., Wyer, M., & Epstein, J. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995. https://doi.org/10.1037/0022-3514.70.5.979.

Dillard, J. P., Segrin, C., & Harden, J. M. (1989). Primary and secondary goals in the production of interpersonal influence messages. *Communication Monographs, 56,* 19–38. https://doi.org/10.1080/03637758909390247.

Dzindolet, M. T., & Pierce, L. G. (2004). *A computerized text analysis can detect deception.* Paper presented at the Human Factors and Ergonomics Society Annual Meeting, New Orleans, LA.

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage.* New York: W.W. Norton.

Elntib, S., Wagstaff, G. F., & Wheatcroft, J. M. (2015). The role of account length in detecting deception in written and orally produced autobiographical accounts using reality monitoring. *Journal of Investigative Psychology and Offender Profiling, 12,* 185–198. https://doi.org/10.1002/jip.1420.

Goffman, E. (1959). *The presentation of self in everyday life.* New York: Anchor Books.

Hancock, J. T., Birnholtz, J. P., Bazarova, N. N., Guillory, J., Amos, B., & Perlin, J. (2009). *Butler lies: Awareness, deception and design.* Paper presented at the Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).

Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review, 19,* 307–342. https://doi.org/10.1177/1088868314556539.

Knapp, M., Daly, J., Albada, K., & Miller, G. (2002). Backgrounds and current trends in the study of interpersonal communication. In M. Knapp & J. Daly (Eds.), *Handbook of interpersonal communication* (pp. 3–20). Thousand Oaks, CA: Sage.

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111,* 8788–8790. https://doi.org/10.1073/pnas.1320040111.

Larcker, D. F., & Zakolyukina, A. A. (2012). Detecting deceptive discussions in conference calls. *Journal of Accounting Research, 50,* 495–540. https://doi.org/10.1111/j.1475-679x.2012.00450.x.

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D., & Gorski, S. (2011). Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. *Journal of Applied Communication Research, 39,* 19–37. https://doi.org/10.1080/00909882.2010.536844.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392. https://doi.org/10.1177/0261927x14535916.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs, 66,* 125–144. https://doi.org/10.1080/03637759909376468.

Margolin, D., & Markowitz, D. M. (2018). A multitheoretical approach to big text data: Comparing expressive and rhetorical logics in Yelp reviews. *Communication Research, 45,* 688–718. https://doi.org/10.1177/0093650217719177.

Markowitz, D. M., & Hancock, J. T. (2014). Linguistic traces of a scientific fraud: The case of Diederik Stapel. *PLoS ONE, 9,* e105937.

Markowitz, D. M., & Hancock, J. T. (2016). Linguistic obfuscation in fraudulent science. *Journal of Language and Social Psychology, 35,* 435–445. https://doi.org/10.1177/0261927x15614605.

Markowitz, D. M., & Hancock, J. T. (2017). The 27 Club: Music lyrics reflect psychological distress. *Communication Reports, 30,* 1–13. https://doi.org/10.1080/08934215.2016.1210663.

Markowitz, D. M., & Hancock, J. T. (2018). Deception in mobile dating conversations. *Journal of Communication, 68,* 547–569.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29,* 665–675. https://doi.org/10.1177/0146167203029005010.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). *Finding deceptive online spam by any stretch of the imagination.* Paper presented at the 49th Annual Meeting of the Association for Computational Linguistics, Portland, OR.

Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us.* London: Bloomsbury Press.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic Inquiry and Word Count. www.liwc.net.

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PLoS ONE, 9,* e115844.

Pfiffner, J. P. (1999). The contemporary presidency: Presidential lies. *Presidential Studies Quarterly, 29,* 903–917. https://doi.org/10.1111/1741-5705.00073.

Rude, S., Gortner, E.-M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition and Emotion, 18,* 1121–1133. https://doi.org/10.1080/02699930441000030.

Schelleman-Offermans, K., & Merckelbach, H. (2010). Fantasy proneness as a confounder of verbal lie detection tools. *Journal of Investigative Psychology and Offender Profiling, 7,* 247–260. https://doi.org/10.1002/jip.121.

Schober, M. F., & Glick, P. J. (2011). Self-deceptive speech: A psycholinguistic view. In C. Piers (Ed.), *Personality and Psychopathology* (pp. 183–200). New York: Springer.

Serota, K. B., & Levine, T. R. (2015). A few prolific liars: Variation in the prevalence of lying. *Journal of Language and Social Psychology*. https://doi.org/10.1177/026 1927x14528804.

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00420.

Toma, C., & Hancock, J. T. (2012). What lies beneath: The linguistic traces of deception in online dating profiles. *Journal of Communication, 62,* 78–97. https://doi.org/10.1111/j.1460-2466.2011.01619.x.

Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology, 11,* 69–89. https://doi.org/10.17161/str.1808.6098.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). West Sussex: Wiley.

Woolley, J. T., & Peters, G. (2009). The American Presidency Project. Retrieved from http://www.presidency.ucsb.edu/.

Verifiability Approach: Applications in Different Judgmental Settings

Galit Nahari

Evidence is a key component in the conviction of a criminal suspect. Police forces invest significant effort in the collection of evidence that links suspects to crimes, while guilty suspects aim to keep the police from finding this evidence. Sometimes, criminals intentionally attempt to avoid leaving traces at the scene or to destroy existing evidence. For example, they may use gloves to avoid leaving fingerprints, burn the body of their victim to prevent detection of their own DNA, destroy a camera before committing their criminal act, or shred a document that may incriminate them. In the interrogation room, guilty suspects are guided by the same motivation—to keep the police from linking them to the crime. For this purpose, suspects must be careful not to provide details that may incriminate them. This motivation of guilty suspects is a basic foundation of the Verifiability Approach (VA; Nahari, Vrij, Fisher, 2014a, 2014b).

In the interview room, suspects provide accounts such as alibi claims (i.e., the suspect argues that at the time the crime took place, he/she was somewhere else and thus could not have committed the crime) or alternative explanations for their presence at the crime scene (i.e., the suspect admits to being at the crime scene at the time the crime occurred, but provides a

G. Nahari (✉)
Department of Criminology, Bar-Ilan University, Ramat Gan, Israel
e-mail: galit.nahari@biu.ac.il

213

reason other than involvement in the crime). While providing their accounts, liars apply certain strategies in an attempt to be convincing (e.g., Hartwig, Granhag, & Strömwall, 2007; Hartwig, Granhag, Strömwall, & Doering, 2010; Nahari, Vrij, & Fisher, 2012; Nahari et al., 2014b). Clearly, the strategies that liars use are based on their beliefs regarding the way that people behave when they tell the truth. Based on the very popular belief that truthful accounts are rich in detail (Bell & Loftus, 1989; Johnson, 2006; Johnson, Foley, Suengas, & Raye, 1988), liars generally wish to provide many details in their accounts (Hartwig et al., 2007; Nahari et al., 2012). However, the provision of details is a risk for liars, because examination of these details may reveal them to be lying. Liars are aware of this danger (see Masip & Herrero, 2013; Nahari et al., 2012) and may thus be inclined to avoid mentioning false details. This *liars' dilemma* hypothesis is the basis of the VA (Nahari et al., 2014a, 2014b).

On one hand, liars are motivated to include many details so that they appear honest, while on the other hand, they are motivated to avoid providing false details, to minimize the chances of being caught. A strategy that serves as a compromise between these conflicting motivations is to provide details that cannot be verified. For example, liars may prefer to describe a car that passed by in the street at a certain moment (difficult to verify) over describing a text chat that they were involved in at a certain point in time (easy to verify). Therefore, according to the VA, when attempting to make an impression of honesty, liars will likely choose to provide details that are difficult to verify and to avoid providing details that are easy to verify. There is empirical support for the VA, which has been found valid in distinguishing truths from lies (see Vrij & Nahari, 2017, 2018). A recent book chapter by Nahari (2018) describes in detail the development of the approach alongside its theoretical framework and components.

Yet, the VA was developed with police interrogation settings in mind, specifically for distinguishing between false and truthful accounts provided by suspects. This must be taken into consideration before applying the approach in other contexts or settings. In the current chapter, I first describe the VA protocol and coding system in relation to its theoretical framing with respect to suspect accounts in the police interrogation setting. I will then discuss key differences between several judgment settings, how these differences influence the application of VA, and which adjustments should be made to maximize the VA's potential when applying it in different settings.

Application of the VA to Assess Suspect Accounts in Police Interrogation Settings

As noted above, the VA was developed for examining the veracity of suspect accounts within the context of police interrogations. Thus, the original protocol of the approach was built to serve this purpose specifically. In the current section, this protocol is described in detail.

VA Protocol for Suspect Accounts

The VA protocol has two phases: collection of statements and coding of statements.

Statement Collection

The VA requires a free-recall account of a specific event. The interviewer must be clear about the beginning and end of the time frame of interest, and ask the interviewee to provide as much detail as possible. The instructions should follow the basic structure of this example: *"You are suspected of … [the specific crime]. Please tell me what you were doing during … [specific time period; e.g., "during last weekend"], from… [specific time, so the suspect knows where to start the account; e.g., "from Friday at 20:00" or "from the moment Fred arrived at the flat"] to… [specific time, so the suspect knows where to end the account; e.g., "to Sunday at 20:00" or "until Fred left the flat"]. When you are ready, please tell me about your activities in as much detail as possible, and do not exclude anything, so I can have an idea of what happened during … [specific time period]. Be sure that you mention all details, activities, people involved, and conversations that took place, etc. Give as much information as you can, including information that seems irrelevant to you."*

Studies have shown that informing interviewees about the mechanism of the VA facilitates the efficacy of the approach in distinguishing lies from truths (e.g., Nahari et al., 2014b). It appears that informed liars do not provide more verifiable details than do uninformed liars, presumably because they do not have truthful verifiable details to provide, while informed truth-tellers provide more verifiable details than uninformed truth-tellers, presumably because they are more aware of the importance of these details, and are also able to provide them. As a result, the difference in level of verifiability between liars and truth-tellers tends to be greater among informed interviewees than among uninformed interviewees. This finding led to the inclusion of information regarding the VA's mechanism as an integral component of the VA protocol.

Interviewees should be informed before providing their statements, using a format similar to this example: *"We are going to review your statement carefully and check whether or not the truthfulness of the details you provided can be verified. We know that liars prefer to avoid details that we can verify, whereas truth-tellers prefer to provide details that we can verify. Therefore, a verifiable statement is more convincing, and the amount of verifiable details you provide in your statement can be critical for you. It is recommended that you include in your statement as many verifiable details as possible."*

While the suspect provides his or her statement, it is very important not to interrupt by asking follow-up questions or commenting. As the VA is a strategy-based technique, the voluntary provision of verifiable details is most relevant.

Statement Coding

In this stage, the statement provided by the interviewee is coded. Specifically, a trained coder identifies and counts the verifiable details in the text (Nahari, 2016, 2018). As interviewees are asked to describe in detail what they were doing during a specific time frame, that is, to recall an episodic memory of a specific event, their statements may include: (a) perceptual details such as things they saw (e.g., a woman walked her dog, there were three free tables, Mike smoked a cigar); actions (e.g., I jumped, we laughed, I pushed the button); sounds (e.g., the radio was on, he told me that the exam was difficult, the phone rang, I was woken up by the alarm clock); tastes (e.g., I had a coffee, I tasted the soup before I served it, the tea was sweet); smells (e.g., we could smell the smoke, I could smell her perfume); (b) contextual details regarding times and locations, for example, dates, seasons, weekdays, hours, duration of time, order of activities, spatial arrangement of people and objects; (c) emotions, describing how the suspect felt during the described event (e.g., I was upset, it disappointed me, my heart jumped with joy); (d) thoughts that crossed the suspect's mind during the described event (e.g., I thought that it was rude behavior, it reminded me that I have an exam next week, I tried to make up my mind whether to join him or not); (e) reasoning (e.g., I went to the principal's office *to complain about the service I got*); and (f) the suspect's inferences and interpretations regarding the event (e.g., I believe I had my watch with me at the pub as I remember that I told someone the time, she was pleased).

Working definition of verifiable details. Put simply, a detail is verifiable if its truthfulness can potentially be checked. From this, one can deduce that emotions, reasoning, inferences, and interpretations are never considered verifiable details, as they are subjective. There are no means to determine the truthfulness of how someone felt or what someone thought during an event. Similarly, inferences and interpretation are subjective opinions on reality. It is possible to argue that someone's inference was wrong (e.g., he mistakenly thought that Fred was disappointed) but it is difficult to prove that he intentionally provided a wrong inference (e.g., he really believed that Fred was disappointed). Thus, only perceptual and contextual details can, potentially, be verifiable.

In accordance with this distinction, the working definition is as follows: *Verifiable details are perceptual and contextual details that are related to occurrences that were (i) documented; (ii) carried out together with (an)other identified person(s); or (iii) witnessed by (an)other identified person(s)* (Nahari et al., 2014a, 2014b). A detail that cannot be related to an occurrence that was documented, carried out, or witnessed by (an)other person(s) is a non-verifiable detail. The three components of the definition are described in further detail below.

Details related to occurrences that were documented. Activities are said to be documented when they leave traces that can subsequently be checked. Documentation can occur either manually (e.g., a name on a waiting list at

a restaurant) or via technology (e.g., closed-circuit television [CCTV]). For more details, see Nahari's (2018) work.

Details related to occurrences that were carried out together with (an) identified person(s). Sometimes interviewees describe activities or events that were carried out with others. If the persons mentioned are identifiable and traceable, they can be approached to verify the truthfulness of the details provided by the interviewee. Consider this fictive statement provided by an interviewee: "…I was sitting next to him on the bench in Hayarkon Park. While watching children play football, we talked about the surprising results of the elections in the US. We had been sitting there for approximately 15 minutes when a woman approached us and asked for directions…." The verifiability of the details provided in this statement depends on the identity of the man mentioned. If he was introduced in the full statement as someone who could be traced (e.g., acquaintance, family member), then the statement includes many verifiable details. In fact, all the activities, times, and locations provided in the statement can be checked by asking the man about them. If, in contrast, he was introduced as a stranger who could not be traced (e.g., a stranger who happened to be at the park, and sat with the interviewee on the bench), none of the provided details are verifiable.

Details related to occurrences that were witnessed by (an) identified person(s). This component of the definition is very similar to the previous one. The only difference is that the other person(s) was (were) not said to have carried out the activities with the interviewee, but only to have witnessed them. To illustrate this, consider an account in which the interviewee described sitting alone in the public library, on a certain day, reading a book for half an hour. The interviewee also mentioned that during this time, two ladies were sitting at a table next to him and could see him the entire time he was there. Now, if the interviewee introduced the two ladies as individuals who could be traced (e.g., colleagues from work), all the details he provided would be verifiable (e.g., his presence in the library, duration of presence, where he sat, that he was reading a book). In contrast, if the two ladies could not be traced (i.e., the interviewee introduced them as strangers), none of these details could be checked.

Working assumptions and principles underlying the coding process. This section addresses assumptions and principles that should be used to guide VA coding.

The reasonable standard person assumption. The VA posits that verifiability level can help determine veracity because liars avoid providing verifiable details as a strategy. Thus, the presence of verifiable details in an account is only significant if the interviewee is aware that these details are verifiable. For this reason, details are identified as verifiable only if it is likely that the common interviewee would be aware of their documentation. In other words, I use the *reasonable person standard* in determining the verifiability of a detail. For example, it can be confidently argued that the reasonable person assumes that the occurrence of a specific phone call is verifiable, because it

is a well-known fact that communication companies document phone calls. Yet, in the case of CCTV, the reasonable person standard is not applicable. It is difficult to predict whether the common interviewee would be aware that CCTV is present at a particular location, as this awareness would be based on familiarity with the specific location rather than on common knowledge or assumptions of plausibility. For example, CCTV is frequently found in bars in England. Yet, whether it is actually present in a specific bar is a matter of familiarity with that bar (rather than a general understanding regarding the presence of CCTV in bars), which is difficult to predict. Thus, alleged activities that could have been caught by the CCTV will be considered verifiable only when the interviewee explicitly mentions the existence of CCTV on the premises.

Event-relevant details. Only details that are related to the relevant event should be coded. The relevant event is defined by the question asked. In our context, the interviewee is asked to describe his/her whereabouts during a certain period of time (e.g., please tell me what you did last Friday, from 8:00 to 14:00). One should avoid coding information that is external to the event under question (in this case, the interviewee's whereabouts on Friday between 8:00 and 14:00), including information that is considered to be a "known fact." For example, while providing their statements, interviewees can mention future events (e.g., tomorrow I have a driving test); past events that occurred outside the relevant time period (e.g., at my first day at work as a technician…, on Friday evening I had guests, on the next day I slept till noon); facts about their routines or habits (e.g., every morning, I eat breakfast while reading the newspaper). In addition, interviewees can provide descriptions of people and objects or indicate facts that they could know independently from the relevant event (e.g., a student who is familiar with the campus buildings, and describes one of them in detail, or a person that mentions how many brothers and sisters he has). These types of information are external to the relevant event and should be ignored in coding.

Avoid over-crediting. Many times, interviewees repeat information. For example, in this statement: "I arrived at the office at 11:00. It took me 30 minutes to finish my business there, so I left at 11:30," there are four contextual details: arriving at the office, time of arrival, leaving the office, time of departure. The duration of time does not add new information (as the interval between the time of arrival and the time of departure can be calculated) and thus does not receive additional points.

THE APPLICABILITY OF THE VA PROTOCOL IN DIFFERENT SETTINGS

It is not automatically the case that a tool developed for a specific context or setting will be applicable in other settings. In fact, the extension of a tool to different applications *must* be accompanied by theoretical adjustments and

empirical examinations. In the current section, I will discuss the applicability of the VA, which was originally developed for examining the veracity of suspect accounts in the police interrogation setting, to several different settings. I will start with veracity testing among victims and witnesses in the police setting, continue with insurance and malingering settings, and conclude with the airport security setting.

Application of the VA for Testing Veracity Among Witnesses and Victims

In addition to suspects, the police may also interview and assess the credibility of witnesses and victims. There is no reason to doubt the applicability of the VA protocol, as described above, to the case of a witness or a victim. In both cases, the interviewee can be asked to describe the relevant event in detail and to include as many verifiable details as possible. The coding system described above is also appropriate here.

An exception is the case of alibi witnesses. An alibi witness is a person who, by means of his/her statement, positions a suspect at a location other than the crime scene (Dahl & Price, 2012). Alibi witnesses are frequently used by defendants in Canadian and US American criminal court cases (Burke & Turtle, 2003). Studies show that people believe they would be able to find a witness to corroborate their false alibis (Culhane, Hosch, & Kehn, 2008), and that many people are indeed willing to lie for others (Hosch, Culhane, Jolly, Chavez, & Shaw, 2011). As such, it is highly relevant to examine the efficiency of existing tools in assessing the veracity of alibi witness accounts.

The primary claim of an alibi witness is generally that "the suspect was with me at the relevant time, and thus has an alibi." Consequently, the alibi witness not only has to prove his/her whereabouts at the relevant time, but also has to show that the suspect was with him/her at the relevant time (see Nahari & Vrij, 2014; Vernham et al., 2018). Thus, the original definition of a verifiable detail, as used with reference to suspect accounts, is slightly different for alibi witness accounts. This is best illustrated by an example, in which an interviewee provided the following statement: "…we went to the ATM at the mall, where I withdrew some money. Then, we had a meeting with my insurance agent, during which we discussed a private health insurance plan for me…" For a suspect interviewee, this statement includes six verifiable details: (i) ATM (contextual; location); (ii) at the mall (contextual; location); (iii) withdrawing money (perceptual; action); (iv) then (contextual; order of activities); (v) meeting with the insurance agent (perceptual); and (vi) discussing the health insurance plan (perceptual). Withdrawing money is a documented activity; it is possible to verify whether it occurred and, if so, at what time and location. It is also possible to verify the occurrence and content of the meeting with the insurance agent, as a third (identifiable) person was involved (i.e., the agent). However, for an alibi witness interviewee, because

of the need to verify the involvement or presence of the suspect as well, the same statement includes only two verifiable details, namely the meeting with the insurance agent and the discussion about the health insurance plan. The insurance agent can say whether the suspect was present at the meeting with the witness. However, it is not possible to verify whether the suspect was present while the witness withdrew money (unless the witness explicitly indicated the presence of CCTV at the ATM). To date, two studies (Nahari et al., 2014a; Vernham et al., 2018) have provided empirical support for the validity of the VA, using the coding method described above, in assessing the veracity of alibi witness accounts.

Application of the VA in the Insurance Setting

An insurance claim is a request presented to an insurance company for payment related to a theft, loss, accident, illness, damage to property, etc. Similar to suspect accounts, descriptions of insurance incidents to justify payment address past events that can be either truthful or false. Thus, it would appear to be possible to generalize the original VA protocol, developed for police interrogation settings, to include the insurance setting without adjustments. In such cases, the claimant would describe the insurance incident carefully and provide as many verifiable details as possible. This description would then be coded, using the working definition of verifiable details given above. However, in spite of the similarity, there is one fundamental difference between the settings that merits attention. The police often know where and when a crime under investigation occurred. As such, they ask interviewees to report what they were doing at a certain time. The case is entirely different in the insurance setting, in which the claimant is the one who tells the police (or the insurer) where and when the incident (loss, theft, or damage) occurred. This allows liars to choose a truthful event and to embed a false insurance incident within this event. For example, a liar can describe a birthday party that he actually attended recently, and tell the police that his phone disappeared during that party. In this case, the liar is able to provide many truthful verifiable details about the party (e.g., who was there, descriptions of guests' clothing, content of conversations, descriptions of food and drinks), while embedding the false claim that his phone disappeared while he was there (see Nahari, Leal, Vrij, Warmelink, & Vernham, 2014).

This difference between the settings most likely explains why the information protocol, which is recommended for police interrogation settings, is crucial in insurance settings (Harvey, Vrij, Nahari, & Ludwig, 2016; Vrij, Nahari, Isitt, & Leal, 2016), as demonstrated by the finding that verifiable details discriminated between liars and truth-tellers when the information protocol was employed (Vrij et al., 2016), but not when it was omitted (Nahari et al., 2014). Harvey et al. (2016) manipulated use of the information protocol, replicating this pattern and further showing that in the insurance setting, the

information protocol not only facilitated the VA, but was critical to its ability to distinguish between true and false claims. Presumably, in the insurance setting, where the integration of truthful details is frequent, differences between liars and truth-tellers are small, and thus require facilitation to be detected. The information protocol, which increases the differences between liars and truth-tellers (Nahari et al., 2014b) serves as a facilitator.

Insurance claims do not always refer to a specific incident such as an accident or burglary. In the case of illness, for example, the claim may include references to a medical diagnosis and physical symptoms instead. If the claimant can show clear medical evidence supporting the existence of illness (e.g., a CT scan showing a tumor, which a doctor confirms is malignant), questions of credibility with respect to the existence of the illness (e.g., cancer) will rarely be raised. Yet, there are illnesses and physical conditions that leave few or no traces that can be observed or documented. For example, in the case of tension headaches, nothing abnormal would be found in the course of a general physical or neurological examination. It is in such situations that malingering becomes an issue. The malingering setting is discussed in the next section.

Application of the VA in the Malingering Setting

Malingering is the fabrication of symptoms of mental or physical disorders with the aim of gaining a secondary benefit, such as financial compensation, avoiding an activity (e.g., attending school, an exam, or a meeting), avoiding or minimizing punishment, or attracting attention. Symptoms like pain, fear, headaches, and fatigue are subjective in nature. Similar to feelings and thoughts, symptoms cannot be observed or documented, unlike external signs like high blood pressure or rash. This makes the application of the VA for detecting malingering difficult. Indeed, the only published paper to date concerning malingering (Boskovic, Bogaard, Merckelbach, Vrij, & Hope, 2017) provides weak evidence for the efficacy of the VA in this setting.

The main challenge in the malingering setting is related to the working definition of verifiable detail. According to the original definition used in VA coding (see above), information is verifiable when there is an indication (e.g., some kind of documentation or witness) confirming its truthfulness. In the police interrogation setting, a piece of information can refer to an activity, location, duration, or presence of an object or a person, among other things, all of which can be observed or documented. In the malingering setting, one is concerned with the fabrication of physical or mental symptoms (rather than activities, locations, times and the like), which cannot be observed or documented. To illustrate the problem, consider that a person claims he was unable to attend an exam due to a stomachache. Obviously, his ache cannot be "shown," as he cannot provide any documentation of it or suggest a witness who can confirm its existence. What he can do is provide secondary-level

indications of the ache, or activities that indirectly support his claim. For example, he might show that he had visited his doctor, bought medication, undergone medical examinations (e.g., blood test), notified his professor that he would not attend the exam for medical reasons, was observed squirming in agony by someone. These secondary-level indications can only *support* the claim regarding the symptom (i.e., stomachache), but cannot confirm the existence of the symptom itself. In this sense, the verifiability of an account in the malingering setting is weaker by definition. In any case, the original working definition of verifiable details is inappropriate for malingering and must be redefined or adjusted in accordance with the unique nature of this setting.

Application of the VA in the Airport Security Setting

Airport security is a particularly challenging setting for deception detection in general, especially because of the need to check many people at the same time (Kleinberg, Nahari, & Verschuere, 2016), the multicultural environment (Jupe, Leal, Vrij, & Nahari, 2017), and the variety of possible hazards (e.g., terror, drug trafficking, illegal immigration) that must be addressed. A primary component of deception in the airport setting is related to purpose of travel. Clearly, from a security perspective, it is highly important to detect travelers with illegal intentions.

The main difference between suspect accounts (in the police interrogation setting) and traveler statements regarding travel plans (in the airport setting) is that the former deal with past events and the latter with future events. In the airport setting, the interviewee can be asked to describe in detail what he/she is going to do in the destination country and to provide as many verifiable details as possible. The original working definition thus can be applied to the airport setting, following an adjusted interpretation of its components. It is possible to check future plans and activities based on documentation and on interviews with identified people. For example, consider an Israeli traveler who claims to be on an academic trip to Vienna. In her statement, she claims that she will be attending a consortium meeting, where she will meet people from Israel and from European countries including Austria, Belgium, Spain, Portugal, and Romania. She adds that she will be staying at the Hotel City Central and that on the final day of her trip, at 11:00 am, she is going to meet her Belgian and Portuguese colleagues together with three Israeli police officers (all partners in the consortium) on the other side of the city. Many verifiable details can be identified in her statement: length of the stay in Vienna (indicated by flight tickets), attendance at a consortium meeting (indicated by meeting registration documents), others who will attend and their nationalities (indicated by meeting registration documents), accommodation details (indicated by hotel booking records), and all the details of the meeting on the final day of the trip (indicated by interviews with other participants).

The single VA study conducted within the airport setting to date (Jupe et al., 2017) supported the validity of the approach in this setting. There are, however, two features of airport security that merit attention. First, accounts regarding future events are expected to be poorer in verifiable details in comparison with accounts regarding past events, because only the latter can include details about unexpected events (e.g., bumping into someone, having an unplanned phone call). Second, travelers with illegal intentions often create a cover story in advance to mislead airport security. A person planning to hijack a flight to London can, for example, book a two-way flight ticket, book a hotel in London, and inform relatives in London that he is coming to visit. It is easier to establish a cover story for a future event as, in contrast to a past event, its accuracy cannot be checked in real time. These two unique features of future events may constitute challenges in applying the VA to the examination of veracity in travel purpose statements.

Summary

In the current chapter, I discussed the applicability of the VA, developed for examining the veracity of suspect accounts in police interviews, in different settings. Beyond the practical aspects of this discussion, I demonstrated the importance of considering the idiosyncratic features of each setting in comparison with the setting in which the approach was developed, and the need to adjust the original protocol in accordance with these features. I began by discussing use of the VA to assess the statements of victims and eyewitnesses in the police interrogation setting, which differ in terms of the perspective presented of the same past criminal event. Next, I discussed the insurance claim setting, in which we also deal with a past event, though different in nature from the criminal one. Third, the malingering setting was addressed, revealing a new issue related to the verifiability of mental or physical symptoms that cannot be observed or documented directly. Finally, factors pertaining to examination of future events were discussed in the context of the airport security setting. While there are undoubtedly additional settings in which the VA may be beneficial (e.g., assessing veracity of identity claims), the focus of this chapter was on settings that have already been studied empirically, as an illustration of the issues that should be considered when attempting to extend the range of the VA.

References

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology, 56,* 669–679. https://doi.org/10.1037/0022-3514.56.5.669.

Boskovic, I., Bogaard, G., Merckelbach, H., Vrij, A., & Hope, L. (2017). The verifiability approach to detection of malingered physical symptoms. *Psychology, Crime & Law, 23*(8), 717–729. https://doi.org/10.1080/1068316X.2017.1302585.

Burke, T. M., & Turtle, J. W. (2003). Alibi evidence in criminal investigations and trials: Psychological and legal factors. *Canadian Journal of Police and Security Services, 3,* 286–294.

Culhane, S. E., Hosch, H. M., & Kehn, A. (2008). Alibi generation: Data from US his-panics and US non-hispanic whites. *Journal of Ethnicity in Criminal Justice, 6,* 177–199. https://doi.org/10.1080/15377930802243395.

Dahl, L. C., & Price, H. L. (2012). He couldn't have done it, he was with me!: The impact of alibi witness age and relationship. *Applied Cognitive Psychology, 26,* 475–481. https://doi.org/10.1002/acp.2821.

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law, 13,* 213–227. https://doi.org/10.1080/10683160600750264.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Doering, N. (2010). Impression and information management: On the strategic self-regulation of innocent and guilty suspects. *The Open Criminology Journal, 3,* 10–16.

Harvey, A. C., Vrij, A., Nahari, G., & Ludwig, K. (2016). Applying the verifiability approach to insurance claims settings: Exploring the effect of the information protocol. *Legal and Criminological Psychology, 22*(1), 47–59. https://doi.org/10.1111/lcrp.12092.

Hosch, H. M., Culhane, S. E., Jolly, K. W., Chavez, R. M., & Shaw, L. H. (2011). Effects of an alibi witness's relationship to the defendant on mock jurors' judgments. *Law and Human Behaviour, 35,* 127–142. https://doi.org/10.1007/s10979-010-9225-5.

Johnson, M. K. (2006). Memory and reality. *American Psychologist, 61,* 760–771.

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117,* 371–376. https://doi.org/10.1037/0096-3445.117.4.371.

Jupe, L. M., Leal, S., Vrij, A., & Nahari, G. (2017). Applying the verifiability approach in an international airport setting. *Psychology, Crime & Law, 23,* 812–825.

Kleinberg, B., Nahari, G., & Verschuere, B. (2016). Using the verifiability of details as a test of deception: A conceptual framework for the automation of the verifiability approach. In *Proceeding of NAACL-HLT 2016* (pp. 18–25).

Masip, J., & Herrero, C. (2013). 'What would you say if you were guilty?' Suspects' strategies during a hypothetical Behavior Analysis Interview concerning a serious crime. *Applied Cognitive Psychology, 27*(1), 60–70. https://doi.org/10.1002/acp.2872.

Nahari, G. (2016). When the long road is the shortcut: A comparison between two coding methods for content-based lie detection tools. *Psychology, Crime and Law, 22,* 1000–1014.

Nahari, G. (2018). The applicability of the verifiability approach to the real world. In P. Rosenfeld (Ed.), *Detecting concealed information and deception: Recent developments* (pp. 329–349). Amsterdam: Elsevier.

Nahari, G., Leal, S., Vrij, A., Warmelink, L., & Vernham, Z. (2014). Did somebody see it? Applying the verifiability approach to insurance claims interviews. *Journal of Investigative Psychology and Offender Profiling, 11,* 237–243. https://doi.org/10.1002/jip.1417.

Nahari, G., & Vrij, A. (2014). Can I borrow your alibi? The applicability of the verifiability approach to the case of an alibi witness. *Journal of Applied Research in Memory and Cognition, 3,* 89–94. https://doi.org/10.1016/j.jarmac.2014.04.005.

Nahari, G., Vrij, A., & Fisher, R. (2012). Does the truth come out in the writing? SCAN as a lie detection tool. *Law and Human Behavior, 36,* 68–76. https://doi.org/10.1007/s10979-011-9264-6.

Nahari, G., Vrij, A., & Fisher, R. P. (2014a). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology, 19*, 227–239. https://doi.org/10.1111/j.2044-8333.2012.02069.x.

Nahari, G., Vrij, A., & Fisher, R. P. (2014b). The verifiability approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology, 28,* 122–128. https://doi.org/10.1002/acp.2974.

Vernham, Z., Vrij, A., Nahari, G., Leal, S., Mann, S., & Satchell, L. (2018). *Applying the verifiability approach to deception detection in alibi witness situations.* Manuscript Submitted for Publication.

Vrij, A., & Nahari, G. (2017). Verbal lie detection. In P. A. Granhag, R. Bull, A. Shaboltas, & E. Dozortseva (Eds.), *Psychology and law in Europe: When West meets East* (pp. 263–282). London: CRC Press.

Vrij, A., & Nahari, G. (2018). The verifiability approach. In J. Dickinson, N. Schreiber Compo, R. Carol, & M. McCauley (Eds.), *Evidence-based investigative interviewing.* London: Routledge.

Vrij, A., Nahari, G., Isitt, R., & Leal, S. (2016). Using the verifiability lie detection approach in an insurance claim setting. *Journal of Investigative Psychology and Offender Profiling, 13*(3), 183–197. https://doi.org/10.1002/jip.1458.

Understanding Lie Detection Biases with the Adaptive Lie Detector Theory (ALIED): A Boundedly Rational Approach

Chris N. H. Street, Jaume Masip and Megan Kenny

People believe others are telling the truth more often than they believe others are lying. This is the truth bias. While some theories have seen the bias as an uncontrollable error or default (e.g., Burgoon & Buller, 1994; Gilbert, 1991; Levine, 2014), the adaptive lie detector theory,[1] or ALIED (Street, 2015), argues that the bias is a result of making an informed and adaptive judgment in a situation where there is little useful information available. The account can also explain why it is that people sometimes show a bias to believe others are *lying*, called the lie bias. This chapter (i) explores how the truth and lie biases can be seen as adaptive, (ii) considers future research streams under an adaptive position, and (iii) makes suggestions for how lie detection research can benefit from shifting focus toward theory development over direct application in the absence of theory.

C. N. H. Street (✉)
Department of Psychology, University of Huddersfield, Huddersfield, UK
e-mail: c.street@hud.ac.uk

J. Masip
Department of Social Psychology and Anthropology,
University of Salamanca, Salamanca, Spain
e-mail: jmasip@usal.es

M. Kenny
Department of Psychology, Sociology and Politics,
Sheffield Hallam University, Sheffield, UK
e-mail: Megan.Kenny@hud.ac.uk

ALIED: Bias as an Adaptive Response

To see the truth or lie bias as an adaptive response, we need to first consider the situation a rater is faced with. People display very few cues to deception (DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007), if any (Levine, 2010). If they produce any cues to deception, they are weak and unreliable (DePaulo et al., 2003; Hartwig & Bond, 2011). Under these circumstances, it is perhaps no surprise that raters are inaccurate.

But why should they be biased in this situation? ALIED argues that raters ordinarily attempt to use reliable indicators that directly and causally relate to the statement being judged. These are referred to as individuating cues because they individuate that statement. For instance, if I claim to have been in France last week, a reliable indicator related directly to that statement would be CCTV footage of my presence in a French store. The use of reliable indicators like these is promoted by the strategic use of evidence technique (Granhag & Hartwig, 2014) and the verifiability approach (Nahari, Vrij, & Fisher, 2014).

Alas, the rater is rarely so fortunate to have this information available. Indicators of deception, such as nonverbal behaviors, are often unreliable (DePaulo et al., 2003). One option for the rater in this situation is to guess randomly, perhaps by flipping a mental coin. Another, more adaptive strategy would be to make use of information that does not directly relate to *this* statement ('I went to France last week'), but rather to statements in general. This sort of information is referred to by ALIED as context-general information. Thankfully, most of us interact with people who tell us the truth most of the time (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Halevy, Shalvi, & Verschuere, 2014). In fact, if language is to be useful as a communicative tool, people need to tell the truth more often than they lie (Grice, 1975; Sperber & Wilson, 1995). In the absence of a reliable individuating cue, raters can use their prior knowledge of the situation ('people tell the truth most of the time') to make an informed 'guess'—a guess of truth. Relying on a context-general guess will lead to a bias toward making truth judgments. In this way, ALIED views the truth bias as an adaptive response (Street, 2015), not an error or default (Gilbert, 1991; Levine, 2014).

In some situations, context-general information suggests people will lie. For instance, the rater may be deciding whether a salesperson's enthusiasm for the used car they are pitching is feigned or genuine. Feigning of this sort is not uncommon in the sales community (see Carr, 1968), which may lead to a 'salespeople tend to lie' context-general belief and in turn a lie bias (DePaulo & DePaulo, 1989). Thus, ALIED proposes that people use reliable individuating cues when they are available, but in their absence, will rely on context-general information to make an informed judgment.

ALIED takes quite an unusual view of both the low accuracy and the bias to believe others. It is an adaptively rational account built on boundedly rational decision-making theories (e.g., Simon, 1990) and can be modelled

as a Bayesian reasoner (Street, Bischof, Vadillo, & Kingstone, 2015). We consider how the field of lie detection can move forward with a more adaptive view of the lie detector by considering a number of potential research streams that could be undertaken. While we use ALIED as our case theory, the approaches could be applied to other theoretical accounts that are concerned with whether raters make adaptive (in the sense of functional and flexible) or error-prone judgments.

Cognitive Bases of Response Biases

One adaptive view of the truth bias, as adopted by ALIED, is that it is a satisficing judgment (see Simon, 1990) based on context-general information in the absence of more reliable information. Put another way, the truth bias is a 'good guess' in the current situation when there are not more reliable clues around. A number of testable predictions regarding the truth bias are outlined in the original theory article (Street, 2015) and so will not be revisited here. This section explores recent tests of the account and introduces new research questions that need to be addressed if an adaptive perspective is to hold water, with a particular focus on the lie bias.

Recent Tests

The ALIED account of the truth bias has been tested in recent work. In contrast to the Spinozan account claim that the truth bias is automatic (Gilbert, 1991), there is evidence that a bias toward truth judgments is a guess under uncertainty (Hasson, Simmons, & Todorov, 2005; Street & Kingstone, 2016; Street & Richardson, 2015a). Whether respondents are truth or lie biased depends on the current context (Nadarevic & Erdfelder, 2013; Street et al., 2015; Street & Kingstone, 2016; Street & Richardson, 2015b; see also Richter, Schroeder, & Wörhmann, 2009).

In the most direct test of ALIED, Street et al. (2015) demonstrated that with highly reliable individuating clues, there is no response bias observed. Only when those individuating clues were lacking was a response bias observed. Importantly, the direction of the bias was dependent on the context. If participants were led to believe that most people would tell the truth in this situation, in the absence of reliable individuating evidence, participants made roughly 65% truth judgments. But if participants believed that most people would lie in this situation, in the absence of reliable evidence, participants made roughly 40% truth judgments. That is, participants showed a truth or lie bias that reflected the current context (see also Blair, 2006; Masip & Herrero, 2017). This is consistent with the prediction that the truth and lie biases are functionally equivalent—i.e., as arising from the same underlying cognitive processes. ALIED claims that differences in response arise from different information that is being used by the same cognitive processes.

Lie Bias

When members of the general public are uncertain about the veracity of a statement, they infer truthfulness because most communications they encounter in their daily lives are truthful. However, unlike the general public, some specific professional groups work in contexts where they deal with deception often. For instance, many law-enforcement officers must regularly interview crime suspects who might lie to them. Attorneys, judges, human intelligence experts, auditors, and personnel selection specialists also face deception regularly. Because for these professional groups[2] the base rate of truthful communications is lower than it is for the general public, ALIED predicts that their truth bias will be weak. Empirical results support these contentions: Several reviews and meta-analyses reveal that while 'lie-detection practitioners' are not more accurate than laypeople in judging veracity, they are indeed less truth biased—or even display a *lie* bias (Alonso, Masip, & Garrido, 2010; Bond & DePaulo, 2006; Meissner & Kassin, 2002; Vrij, 2008).

However, some may wonder whether the professional's lie bias is caused by the specific mechanism posited by ALIED. According to ALIED, absent of individuating cues, observers turn to context-general information to judge veracity. But in several experiments, police participants have listed a number of individuating cues (verbal and nonverbal behaviors of the sender) as a basis for their judgments (e.g., Mann, Vrij, & Bull, 2004; Masip, Garrido, Herrero, Antón, & Alonso, 2006; Meissner & Kassin, 2002; Nahari, 2012; Porter, Woodworth, & Birt, 2000). Furthermore, the kinds of cues reported are related to accuracy (Mann et al., 2004) or judgment type (truth vs. lie judgments; Masip et al., 2006; Nahari, 2012). These findings suggest that officers do pay attention to and use available behavioral cues when making a lie–truth judgment.

This is not to say that ALIED cannot explain the lie bias. People may be unaware that most behavioral cues are objectively unrelated to veracity and hence may use these cues rather than using context-general information. In fact, people worldwide have strong beliefs about deception cues (Global Deception Research Team, 2006),[3] and practitioners have the same beliefs as laypeople (Strömwall, Granhag, & Hartwig, 2004). Compared to laypeople, practitioners can display a stronger tendency to use unreliable individuating behavioral cues thought to be indicative of lying. Indeed, Street et al. (2015) found that laypeople made around 80% *lie* judgments when they were using individuating cues that they *believed* were reliable indicators of deception (see also Masip, Alonso, Garrido, & Herrero, 2009). So it is important to bear in mind that while ALIED claims the truth bias ordinarily arises from the use of context-general information, if there are perceived to be reliable indicators of deception present, then ALIED would also predict that raters would judge statements as lies.

Biased Information Processing

Deception judges appear to process information in a biased way. In the studies of Masip et al. (2006) and Nahari (2012), both police officers and laypeople judged the veracity of the same statements. However, officers made more deception (and less truthfulness) judgments than laypeople. Importantly, the officers also reported having used more cues associated with deception judgments (and fewer cues associated with truthfulness judgments) than laypeople. What is more, ambiguous cues were interpreted as *lie* indicators by officers and as *truth* indicators by laypeople. These findings led the researchers to suggest that officers had the a priori belief that senders were going to lie and used confirmatory strategies during the veracity assessment task.

Consistent with ALIED, it has been suggested that officers' a priori belief can result from their frequent involvement in deceptive interactions, as well as of their progressive endorsement through their career of a so-called police culture that involves suspicion and cynicism (e.g., Paoline, 2003). Over the years, both these factors might engender among officers a predisposition to believe that the messages produced by others are deceptive (see Hurst & Oswald, 2012; Masip, 2014; Masip, Alonso, Garrido, & Antón, 2005; Masip, Alonso, Herrero, & Garrido, 2016). However, the data do not unambiguously support these contentions. For example, in Masip et al.'s (2016) study, novice officers were as lie biased as veteran officers (conversely, non-officers displayed a truth bias). Apparently, even though novice officers were not as internally distrustful and suspicious as their more seasoned counterparts (see Masip et al., 2016), they mirrored prototypical police behavior (such as making many deception judgments) in police-related contexts.

Note, however, that Masip et al.'s (2006) and Nahari's (2012) studies are fraught with a 'chicken-and-egg' problem. Clearly, it is possible that their police participants did pay attention to behavioral cues and, based on the processing of such cues, concluded that the senders were lying most of the time. But it is also possible that, instead, they first made a deception judgment and then, subsequently, came up with *post hoc* rationalizations for that judgment (see Festinger, 1957; Levine, Asada, & Park, 2006; Nahari, 2017). If so, then the cues listed by respondents might not have been used at all—they may reflect mere after-the-fact explanations for lie judgments made from context-general information, or from automatic or default judgments. It is possible that the officers' heightened concern with deception, derived either from their frequent involvement in deceptive interactions or their perception of lie detection as a police-relevant activity, leads them to make frequent default deception judgments. Note that this possibility is in line with ALIED (and inconsistent with a truth-default account), as both personal (or professional) history with deception and police relevance of the task at hand are modalities of context-general information.

A recent study by Masip and Herrero (2017) sought to examine the police lie bias in an almost-cueless context. The authors wrote several vignettes where the protagonist denied their involvement in a misdeed. Only two pieces of evidence were available, one suggesting that the protagonist was lying and one suggesting

that she/he was telling the truth. For half the vignettes the misdeed involved a crime, while for the other half it involved a non-criminal event. College students, novice officers, and experienced officers read the vignettes and judged whether the protagonist was lying or telling the truth in denying the misdeed. In support of ALIED's contention that the bias is context-dependent, officers made fewer truth judgments than students in judging *crime* vignettes, but no significant difference emerged between the groups in judging *non-crime* vignettes.

Research Directions Under an Adaptive Paradigm

The research reviewed leads to a number of outstanding questions that can be tackled from an adaptive perspective. We hope to encourage lie detection researchers to explore cognitive theoretical accounts by posing a number of research streams. This section tackles the development of beliefs around cues to deception, searching for cues, cue use in the judgment process, tackling decisional uncertainty, reducing bias, increasing accuracy, and ALIED predictions related to the act of deception.

Developing Beliefs

Very little is known about the process of forming a lie or truth judgment. The Spinozan account (Gilbert, 1991) provided a somewhat detailed explanation of the process, although it has not stood the test of time (e.g., Hasson et al., 2005; Mayo, Schul, & Burnstein, 2004; Nadarevic & Erdfelder, 2013; Street & Kingstone, 2016; although see Mandelbaum, 2014, for a defense of the position). Since then, we have seen relatively little theoretical work on the decision process—at least, until recently (Levine, 2014; Street, 2015). ALIED theory offers a high-level description of the decision process and offers novel, testable predictions. Some of these predictions have been developed into computational simulations (Street et al., 2015). But because ALIED is a high-level description, there is much that has gone unspecified. Some still-untested predictions have been described elsewhere (see Section 8 of Street, 2015).

We just made mention of novel predictions. By novel, we mean a prediction that has emerged from the theory but has yet to be directly tested. It is only by providing predictions that have yet to be tested that an account can be falsified. Einstein's theory of gravitation led to the prediction that light is attracted by heavy objects. This could be tested by observing a light shift as a result of a star's light passing close to the Sun, giving the appearance that the star is further away from the Sun than it actually is. This prediction was novel inasmuch as it had emerged as a natural result of the claims of the theory and was as yet untested (Bailey, 2002). Somewhat less grand but more on-topic, a number of untested predictions emerging from ALIED theory are given in the remainder of this chapter and in the original theory paper. Of course, we do not deny that being able to account for past data is important. But fitting a theory to past data alone means that the only tests of the theory are tests

that have already been conducted and have a generally accepted answer, leaving the theory unfalsifiable and untestable.

How do people come to believe which individuating cues, if any, distinguish deception from honesty, or come to believe that a given source of context-general information suggests people in general will lie or tell the truth? Presumably, socialization processes during childhood, people's observations of behavior and apparently credible media sources inform people's beliefs (Global Deception Research Team, 2006; Hurley, Griffin, & Stefanone, 2014; Strömwall et al., 2004). While younger children lack world experience that may help them detect lies (Lee, Cameron, Doucette, & Talwar, 2002), they nonetheless make lie–truth judgments that are arguably rational. For instance, children trust others more when a game incentivizes cooperation rather than competition (Reyes-Jaquez & Echols, 2015) and trust people who have been previously trustworthy or accurate (Brosseau-Liard & Birch, 2011; Brosseau-Liard, Cassels, & Birch, 2014; Clément, Koenig, & Harris, 2004). Children choose to ignore those who have generally been reliable if the child already has reliable individuating information, e.g., from direct observation of the true state of affairs (Brosseau-Liard et al., 2014; Clément et al., 2004). Looking at adults, we see that wrong beliefs about deception are self-reported (Global Deception Research Team, 2006; Strömwall et al., 2004), although at the point of making a judgment they do seem to rely on more accurate beliefs about deception cues (Hartwig & Bond, 2011).

Are context-general beliefs developed in situations where feedback is available? Are people less inclined to trust their own individuating knowledge in favor of those who they perceive as more accurate lie detectors than themselves? Is theory of mind, an ability to think about other people's mental states (Flavell, 2004), employed to understand others' motivations for lying or telling the truth? These are just some of the issues that may be tackled in an adaptive approach to understanding the development of an adaptive lie detector.

Of course, development extends beyond childhood. Diving into a new world will require learning and adjustment. Police officers and prisoners have been found to exhibit a lie bias (Bond, Malloy, Arias, Nunn & Thompson, 2005; Meissner & Kassin, 2002). The second author has explored how judgment biases change and develop as individuals enter the policing community. One stream of research we can envisage involves an exploration of lie–truth judgment formation and how that process adapts in new environments. ALIED theory would predict that a shift in bias should be observed that can be modelled with a simple *naïve Bayes* equation (see Street et al., 2015, equation 1).

Information Search

Being adaptive does not necessarily mean being highly accurate (Jussim, 2012; Simon, 1990). If raters have the wrong beliefs about clues to deception, for instance, then we may expect those wrong beliefs to hinder the

otherwise adaptive judgment process. By way of analogy, writing inaccurate programming code will make an otherwise robust processing system (a computer) fail—'garbage in, garbage out'. When detecting deception, it would appear that practitioners can hold quite strong beliefs about the wrong cues to deception (Masip & Garrido, 2001), and police officers are more confident in their judgments despite not evidencing higher accuracy than laypeople (Meissner & Kassin, 2002). Garrido, Masip, and Herrero (2004) discuss some specific mechanisms (e.g., a lack of disconfirmatory feedback) to explain why officers may hold such strong opinions. But here we wish to emphasize that holding a strong belief in one's own judgment can lead to a selective exposure, showing a preference for selecting information consistent with one's own view (Festinger, 1957; Fischer, 2011). The information being attended to may not be robust, but that should not be taken to suggest that the judgment process itself is ineffective.

To explore selective exposure, one could generate a set of intuitive and sensible hypotheses. But we would encourage researchers to adopt existing theory to generate predictions. Fischer's (2011) model suggests that lie detection practitioners should be more susceptible to selective exposure, because it is those people who are more confident that are more likely to be selective. One might also make predictions about lay raters. If the presence of highly diagnostic individuating cues increases lay raters' confidence in their judgment, then Fischer's model would predict that raters would be inclined to process information that is consistent with the highly diagnostic individuating cues and ignore information that is inconsistent. It may seem unusual to predict that with greater cue diagnosticity comes selective exposure to information, but there is a precedent. The focal account claims that indirect raters can achieve higher accuracy by attending to a smaller number of highly diagnostic cues to deception (Street & Richardson, 2015c; see also Sporer, Masip, & Cramer, 2014), although note that they did not capture raters' confidence.

This leads into another issue that ALIED is currently unable to account for. How do raters use multiple individuating cues to deception? A compensatory explanation would say that raters use multiple cues and weight them by how diagnostic they are. A non-compensatory approach would say that raters select one cue, possibly the most diagnostic, and ignore the others (see Street, 2015; Street et al., 2015; see Pohl, 2011, for a discussion of the issue outside of lie detection). An interesting possibility arising from Fischer's (2011) model is that non-compensatory approaches are used when confidence is high—that is, raters select to expose themselves to limited information and ignore other information that could be used in the judgment. But as confidence decreases, a compensatory approach (that incorporates both decision-consistent and decision-inconsistent information, i.e., not engaging in selective exposure) may be used. In support of this suggestion, there is some evidence showing that lie detectors integrate multiple individuating cues when they are diagnostic (Street & Richardson, 2015c), although non-compensatory take-one-best-cue models can also fit the data (Street et al., 2015).

When individuating cues are directly contradictory, there is some albeit limited evidence suggesting that individuals may ignore them in favor of context-general information (Masip & Herrero, 2017).

Aside from the decision-making aspects of information search, there are also attentional concerns. We can only achieve perfect accuracy at detecting Pinocchio's lies if (i) we are aware that his nose growth is a diagnostic indicator of deception, (ii) if we attend to it, and (iii) if we ignore other less-diagnostic information that may incorrectly suggest honesty. A relatively recent meta-analysis has found that people do attend to the more diagnostic individuating information that is available (Hartwig & Bond, 2011). But it is not clear how visual and verbal information is gathered over the course of a statement, how the trajectory of the decision changes as raters gather new information, and how verbal and visual sources are integrated (if they are). Eye tracking is a tool that gives a good indication of what people are thinking with high temporal resolution, making it well suited to exploring how visual information is obtained from the speaker and how this affects the developing and final judgment.

Reducing Bias

ALIED claims the response bias can be caused by making judgments using context-based information to fill in the gap of more reliable individuating information. All else being equal, preventing context-based judgments should reduce biased responding. In this section, then, discussion focuses on how to encourage participants to 'decide not to decide' in the absence of diagnostic individuating cues, or at least to consider the alternatives.

When investigating suspected child neglect, some have suggested that social workers should understand they are working in a low-information environment and maintain a 'respectful uncertainty' about the parents' intentions, rather than accepting at face value the claims of the parents (Laming, 2009). How might such a respectful uncertainty be encouraged? ALIED has nothing to say about how best to encourage uncertainty and prevent guessing, but some suggestions are offered here that would be consistent with ALIED's claims.

In experiments, participants are typically forced to make either a lie or truth judgment, which might lead them to guess. An approach to reduce bias in such a context was explored by Blair (2006). Participants were trained to detect deception. At the end of the training, one group was told that the training could have biased them toward making lie judgments (Meissner & Kassin, 2002; Masip et al., 2009), and that if they were uncertain about truthfulness, then they should make truth rather than lie judgments. This procedure decreased their lie bias (Blair, 2006).

But a more direct way of reducing response bias would be to give deception raters the option to not guess. After all, in real-world settings one might choose to withhold judgment until more information becomes available. In the limited number of studies where they are given the opportunity to

withhold judgment, there has been some evidence of a reduction in response biases (Street & Kingstone, 2016; Street & Richardson, 2015a).

What encourages people to make a guess under uncertainty? Fox and Tversky (1995) found that when information is ambiguous or unclear, people are more willing to commit to a judgment if they feel confident or knowledgeable in the domain (see also Heath & Tversky, 1991). Police officers feel more confident in their judgments (Meissner & Kassin, 2002), and they also feel more knowledgeable than laypeople in the area of lie detection (Garrido et al., 2004). This has the potential to lead officers to give too much weighting to low-diagnostic individuating cues, or to put too much faith in their context-general knowledge when making a judgment in the absence of diagnostic individuating cues. One avenue to consider is whether raters can be made to feel less confident in their decisions (e.g., by means of false feedback saying that they are making consistently bad decisions) or made to feel less knowledgeable (e.g., by encouraging the rater to compare their judgments against another person who is believed to be more knowledgeable than themselves: see Fox & Tversky, 1995).

Rather than encouraging uncertainty, one may wish to encourage active deliberation of both possibilities—that the speaker could be lying and could be telling the truth (see O'Brien, 2009, for a similar strategy to reduce confirmation bias in criminal investigations). Information in support of a judgment can vary in how difficult it feels to process it. If the information feels cognitively easy to process, then it is more likely to be considered valid and true, and as a result can sway the judgment (Schwarz, 2015)—even if that information is not actually valid (Schwarz, Newman, & Leach, 2016). Similarly, when it feels difficult to *generate* information in support of a judgment, it guides raters to believe that the judgment may be incorrect (Ask, Greifeneder, & Reinhard, 2012). One possibility, then, is to reduce processing ease both for information that supports the suggestion that the speaker is lying and information that suggests truth-telling, e.g., by having raters attempt to list many cues to honesty and deception (Ask et al., 2012). Another possibility is to increase the processing ease of both the possibility of deception and honesty, e.g., by having raters consider only one reason why the speaker may be lying or telling the truth (Ask et al., 2012). Such a situation may encourage more deliberative processing of the possibilities of honesty and deception.[4]

Concerns for being evaluated also encourage people to consider other interpretations of the same information (Tetlock, 1983; Lerner & Tetlock, 1999; see also Chow & Sarin, 2002), as well as engage in more thorough reflection of their decision-making (Lerner & Tetlock, 1999). Tetlock (1983) found that raters were more likely to be more critical of their own judgment and consider alternative possibilities in anticipation of the evaluator's counterargument. As an example, in the context of lie detection, a police officer may need to explain to their superior why they judged a speaker as lying, but the officer does not know whether the superior believes the speaker to

be lying or telling the truth. To defend the judgment, the officer will need to consider the possible arguments that the superior will make in order to effectively counterargue their position. In a review of accountability effects in decision-making (Lerner & Tetlock, 1999), it was concluded that people are more critical of their own judgments in these situations, provided the superior has a legitimate reason for making the rater accountable, and that the superior is sufficiently knowledgeable in the area. They also list other factors that can affect self-critical thinking, but we do not have the space to consider them all here.

We realize that this is a section filled with 'ifs' and 'maybes'. This is in part because of a substantial problem with ALIED theory: It is not sufficiently detailed to explain how guessing occurs or how it may be prevented. But the potential is there for borrowing from other areas of research in order to explore bias reduction as an attempt to prevent or mitigate an adaptive 'guess' under uncertainty.

Making Good Judgments

Professional groups do not outperform non-professionals in terms of detection accuracy (Aamodt & Custer, 2006; Alonso et al., 2010; Bond & DePaulo, 2006; Bradford & Goodman-Delahunty, 2008; King & Dunn, 2010; Meissner & Kassin, 2002; for some explanations, see Cassidy & Buede, 2009, and Vrij, 2004). Similar null effects of expertise on performance have been found in other professional domains (Cassidy & Buede, 2009; Garb, 1989; Lilienfeld, Lynn, Ruscio, & Beyerstein, 2010, Chapter 11).

But skill may be explored in other ways. Before discussing this, we feel it is important to explicitly note that when we talk of 'skill', we are not suggesting that there are individuals who may be more accurate than others (see Bond & DePaulo, 2008). A person may use better strategies than others. For instance, a skilled rater may be able to make judgments at the same (relatively low) rate of accuracy as others but may do so more efficiently. They may be faster in their judgments or need less information, for instance, while maintaining the same level of accuracy as others. That knowledge may be lacking does not mean that the system processing the information is inadequate. There are markers of skill aside from accuracy, and it is these we are considering here.

After observing a set of speakers, and having been provided feedback about their honesty, are some raters better able to deduce which behaviors (if any) are diagnostic and use them in future judgments, or are there situational manipulations that can affect such implicit[5] learning? Of course, cues to deception are typically unreliable and weak (DePaulo et al., 2003), but in situations where there exist highly diagnostic cues (e.g., in Bond, Howard, Hutchison, & Masip, 2013; Nahari et al., 2014), might some individuals be able to make better use of those cues? Do some raters make more consistent

ratings given the same or similar individuating cues and/or context-general information? Are some better able to distinguish between subtle differences in contexts and make use of that information in the absence of more diagnostic individuating cues? And what factors affect the rate at which people incorporate feedback about their judgments? To date, discussion around expertise in the lie detection literature has focused solely on whether there are lie detection 'wizards' who have exceptional, inexplicable accuracy (see Bond, 2008). There are good reasons to be skeptical of an accuracy effect (Bond & DePaulo, 2008). We would suggest instead that the issue of expertise (or more simply, individual difference) would be better focused upon response consistency, ability to discriminate information, and learning efficiency. Unfortunately, much of this work would require a highly controlled stimulus set where the behaviors being displayed by the speakers are known and fully coded.

It should also be borne in mind that raters rarely have access to highly diagnostic individuating cues. ALIED dovetails with the current trend of seeking out techniques that actively create cues to deception, rather than passively hoping to detect them. We will not discuss this here, but instead refer readers to Street (2015). One may also look to develop methods of *preventing* deception in a similar vein. For instance, if you have ever claimed expenses for a work trip, your company likely asked that you submit receipts. These receipts are a necessary, diagnostic, and individuating cue that (we suspect) could reduce the tendency to be deceptive on expense claims.

Lying Ability

In our final section on ALIED, we note that the account can also make testable claims as to what makes for a successful liar. If people use reliable individuating cues when they are available, then a successful liar is one who can generate a reliable cue to honesty and then exploit it at the right moment. In a poker match, a player might choose to scratch their chin whenever they have a potentially winning set of cards and to always reveal their cards at the end of the round. If they have a poor hand, they always fold or leave the game without scratching their chin. In doing so, the player generates a good cue to honesty and provides the opponent with ground truth information. The player could also generate a reliable cue to deception in a similar way. At an opportune moment, when there is a big potential monetary return, the player may have a poor hand. Now is the time for the player to scratch their chin. The prediction from ALIED is that raters will use the cue in their judgment. Therefore, all else being equal and provided the opponent has picked up on the cue, raters will infer honesty and in turn a potentially winning hand. Instead of generating beliefs about individuating cues such as chin scratching, a liar could attempt to harness the beliefs others already have about honesty such as eye contact. In either case, ALIED would predict that people will be consistently below chance accuracy.

In the absence of diagnostic individuating cues, ALIED argues that people use context-general information to fill in the information gap. There is typically more than one piece of context-general information, and these may contradict. For instance, while one may believe that people typically lie, one may also believe that academics typically tell the truth. How then does a rater deal with this conflicting information? One possibility is that information made salient may lead to that information being used more heavily in the judgment (Platzer & Bröder, 2012; Street & Richardson, 2015c). If so, a liar may be able to develop a persuasive argument or use some other means to bring to the attention of the other person a context-general cue that is indicative of honesty. For instance, an unfaithful partner who fears not being believed by their loved one may recall occasions when trust between them was required. Or a lawyer wanting to cast suspicion onto a defendant's alibi may bring jurors' attention to the courtroom as a place where only guilty people find themselves.

Moving Forward

One point we hope to have highlighted in our above discussion is that a relatively simple but clearly defined theory can generate predictions and raise interesting testable questions. While in this chapter we defend an adaptive and boundedly rational view of lie detection, we would actively encourage and welcome clearly defined theoretical discussion, even if they contradict an adaptive perspective. The Spinozan account (Gilbert, 1991) is a good example—its predictions can and have been tested since its inception. And while there is now evidence that is difficult to fit with the theory's claims, it has nonetheless spurred an industry of research and further development. Indeed, it was the starting point of the first author's PhD thesis and it ultimately led to the development ALIED theory, despite ALIED being diametrically situated to the Spinozan account. The testability of a well-defined theory outweighs its predictive success (see Blandón-Gitlin, Fenn, Masip, & Yoo, 2014; Street & Vadillo, 2016, for some brief discussion in the area of lie detection).

It is commonplace in other areas of Psychology to openly share stimuli. In an era of open science and a replication crisis (Open Science Collaboration, 2015), it is becoming all the more important to share resources. Deception researchers tend not to share their stimuli. And yet it is perhaps more important in our field than in most to do so. As researchers, we wish to measure responses to natural forms of deception. That means that not only do our manipulations differ between experiments, but so do our stimuli. Isolating the source of the effect is difficult because stimulus sets are difficult to compare in terms of availability of information. While clues to deception are unreliable and weak (DePaulo et al., 2003), they may vary between different stimulus sets, and across different types of lies (DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007).

If we are to develop theories with tightly defined predictions, it will be necessary to share stimuli in order to replicate research and develop on past findings. Those who are sharing their stimuli will benefit from regular citations and from an understanding that they are developing a stream of research where the stimulus set is not being manipulated between studies, providing greater control. At the same time, conclusions we reach from the research in our own laboratories will necessarily be constrained to the stimulus set we use. It is important to assess generalizability, which will require the use of more than one stimulus set. The Miami University deception detection database (Lloyd et al., 2019) has provided a set of 320 videos of people lying and telling the truth. The stimulus set has been coded and transcribed and can be obtained from http://hdl.handle.net/2374.MIA/6067.

Conclusion

There is a troubling trend in our field at the moment, which the first two authors admit to contributing toward. And that is the lack of theoretical progress. We observe effects, tag an explanation onto them, and hold it as a theory. Our hope with this chapter is to encourage a shift toward a more robust theoretical perspective. Theory should be developed first, from which untested predictions should emerge. It is only by understanding the theory and stating its postulates that we can develop robust practical interventions that will stand the test of time (see Blandón-Gitlin et al., 2014; Blandón-Gitlin, López, Masip, & Fenn, 2017; Sporer, 2016; Street & Vadillo, 2016, for discussions of theory in lie detection work). A theory changes and shifts over time, but can maintain its core. In doing so, it develops original predictions that have yet to be explored. It is the ability to generate clearly defined and novel (untested) predictions that is currently lacking.

Newton's theory of gravitation led to the prediction of Uranus' existence before we had telescopes powerful enough to observe it, in spite of certain oddities (see Bamford, 1996; Kuhn, 1962a). While we have now effectively replaced Newtonian with Einsteinian physics (see Kuhn, 1962b), it would be difficult for anyone to say that Newton's work was not successful.[6] His work has achieved such success because of its predictive validity. A good theory, right or wrong, develops clear and testable predictions. The hope is that this chapter has shown how one can take a relatively simple account and develop as-yet untested predictions from it. We would hope to see authors stop applying theory after-the-fact and to then accept it with little further exploration. While a data-driven account is a good place to start, it should not be the end of testing that account with *novel* predictions that naturally emerge from it. We encourage researchers to not only attach theories to their findings in the discussion, but to offer novel predictions that would come out of it—predictions that are not an attempt to fit existing data, but instead are questions that as yet do not have (much) data trying to answer them. It may be that

the theory will not stand the test of time. ALIED surely won't—it is too simple an explanation to capture the complexities of how people form this social judgment from uncertain information. But because it opens itself to testing, it allows us to progress.

Notes

1. ALIED is not so much a theory as it is a hypothesis. At the time of its publication, it was competing with other hypotheses in the area which were branded as theories, and so the word 'theory' was adopted here too.
2. The professional group having been tested most often by researchers is that of law-enforcement officers. Unless indicated otherwise, the findings reported in the following pages are derived from this group.
3. People verbally respond with poor cues when asked about good cues to deception. That should not necessarily be taken as evidence that people hold wrong beliefs. Indeed, there is evidence to suggest people do use the more reliable cues available (Hartwig & Bond, 2011). We also caution against claims that this must reflect a conscious–unconscious distinction (see Shanks & St John, 1994; Street & Vadillo, 2016). Reasons for the discrepancy include (but are not limited to) that the question is poorly formulated to assess respondents' knowledge, that respondents gave answers that they believed the experimenter expected, and that respondents provided a list of behaviors that they thought other people tend to use, rather than themselves (Strömwall et al., 2004).
4. An alternative possibility is that increasing the processing difficulty will lead to raters disengaging with the task altogether and responding in an unmotivated and less deliberate fashion.
5. The term 'implicit' is sometimes used in the lie detection literature to be synonymous with 'unconscious'. We use the term 'implicit learning' as used in memory research to refer to learning without explicit feedback.
6. To be clear, we are not at all suggesting that our work reaches the lofty heights attained by Newton or Einstein!

References

Aamodt, M., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 16,* 6–11.

Alonso, H., Masip, J., & Garrido, E. (2010). La capacidad de los policías para detectar mentiras [Police officers' ability to detect lies]. *Revista de Derecho Penal y Criminología, 2,* 159–196.

Ask, K., Greifeneder, R., & Reinhard, M. A. (2012). On the ease of (dis)believing: The role of accessibility experiences in credibility judgments. *Applied Cognitive Psychology, 26,* 779–784. https://doi.org/10.1002/acp.2859.

Bailey, A. (2002). *First philosophy: Fundamental problems and readings in philosophy.* Lancashire, UK: Broadview Press.

Bamford, G. (1996). Popper and his commentators on the discovery of Neptune: A close shave for the law of gravitation? *Studies in History and Philosophy of Science Part A, 27,* 207–232. https://doi.org/10.1016/0039-3681(95)00045-3.

Blair, J. P. (2006). From the field: Can detection of deception response bias be manipulated? *Journal of Crime and Justice, 29,* 141–152. https://doi.org/10.1080/0735648X.2006.9721652.

Blandón-Gitlin, I., Fenn, E., Masip, J., & Yoo, A. (2014). Cognitive-load approaches to detect deception: Searching for cognitive mechanisms. *Trends in Cognitive Sciences, 18,* 441–444. https://doi.org/10.1016/j.tics.2014.05.004.

Blandón-Gitlin, I., López, R. M., Masip, J., & Fenn, E. (2017). Cognición, emoción y mentira: Implicaciones para detector el engaño [Cognition, emotion and lying: Implications to detect deception]. *Anuario de Psicología Jurídica, 27,* 95–106. https://doi.org/10.1016/j.apj.2017.02.004.

Bond, C. F., Jr. (2008). A few can catch a liar, sometimes: Comments on Ekman and O'Sullivan (1991), as well as Ekman, O'Sullivan, and Frank (1999). *Applied Cognitive Psychology, 22,* 1298–1300. https://doi.org/10.1002/acp.1475.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Bond, C. F., Jr., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134,* 447–492. https://doi.org/10.1037/0033-2909.134.4.477.

Bond, C. F., Jr., Howard, A. R., Hutchison, J. L., & Masip, J. (2013). Overlooking the obvious: Incentives to lie. *Basic and Applied Social Psychology, 35,* 212–221. https://doi.org/10.1080/01973533.2013.764302.

Bond, G. D., Malloy, D. M., Arias, E. A., Nunn, S. N., & Thompson, L. A. (2005). Lie-biased decision making in prison. *Communication Reports, 18,* 9–19. https://doi.org/10.1080/08934210500084180.

Bradford, D., & Goodman-Delahunty, J. (2008). Detecting deception in police investigations: Implications for false confessions. *Psychiatry, Psychology and Law, 15,* 105–118. https://doi.org/10.1080/13218710701873932.

Brosseau-Liard, P., & Birch, S. A. J. (2011). Epistemic states and traits: Preschoolers appreciate the differential informativeness of situation-specific and person-specific cues to knowledge. *Child Development, 82,* 1788–1796. https://doi.org/10.1111/j.1467-8624.2011.01662.x.

Brosseau-Liard, P., Cassels, T., & Birch, S. A. J. (2014). You seem certain but you were wrong before: Developmental change in preschoolers' relative trust in accurate versus confident speakers. *PLoS ONE, 9*(9), e108308. https://doi.org/10.1371/journal.pone.0108308.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18,* 155–184. https://doi.org/10.1007/BF02170076.

Carr, A. Z. (1968). Is business bluffing ethical? *Harvard Business Review, 46,* 143–153.

Cassidy, M. F., & Buede, D. (2009). Does the accuracy of expert judgment comply with common sense: *caveat emptor. Management Decision, 47,* 454–469. https://doi.org/10.1108/00251740910946714.

Chow, C. C., & Sarin, R. K. (2002). Known, unknown, and unknowable uncertainties. *Theory and Decision, 52,* 127–138. https://doi.org/10.1023/A:1015544715608.

Clément, F., Koenig, M., & Harris, P. (2004). The ontogenesis of trust. *Mind and Language, 19,* 360–379. https://doi.org/10.1111/j.0268-1064.2004.00263.x.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995. https://doi.org/10.1037/0022-3514.70.5.979.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

DePaulo, P. J., & DePaulo, B. M. (1989). Can deception by salespersons and customers be detected through nonverbal behavioral cues? *Journal of Applied Social Psychology, 19,* 1552–1577. https://doi.org/10.1111/j.1559-1816.1989.tb01463.x.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Fischer, P. (2011). Selective exposure, decision uncertainty, and cognitive economy: A new theoretical perspective on confirmatory information search. *Social and Personality Psychology Compass, 5,* 751–762. https://doi.org/10.1111/j.1751-9004.2011.00386.x.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50*(3), 274–290. https://doi.org/10.1353/mpq.2004.0018.

Fox, C. R., & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics, 110,* 585–603.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin, 105,* 387–396. https://doi.org/10.1037/0033-2909.105.3.387.

Garrido, E., Masip, J., & Herrero, C. (2004). Police officers' credibility judgments: Accuracy and estimated ability. *International Journal of Psychology, 39,* 254–275. https://doi.org/10.1080/00207590344000411.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107–119. https://doi.org/10.1037/0003-066X.46.2.107.

Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74. https://doi.org/10.1177/0022022105282295.

Granhag, P. A., & Hartwig, M. (2014). The strategic use of evidence technique. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches* (pp. 231–251). Chichester, UK: Wiley. https://doi.org/10.1002/9781118510001.ch10.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York, NY: Academic Press.

Halevy, R., Shalvi, S., & Verschuere, B. (2014). Being honest about dishonesty: Correlating self-reports and actual lying. *Human Communication Research, 40,* 54–72. https://doi.org/10.1111/hcre.12019.

Hartwig, M., & Bond, C. F., Jr. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*(643), 659. https://doi.org/10.1037/a0023589.

Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological Science, 16,* 566–571. https://doi.org/10.1111/j.0956-7976.2005.01576.x.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty, 4,* 5–28. https://doi.org/10.1007/BF00057884.

Hurley, C. M., Griffin, D. J., & Stefanone, M. A. (2014). Who told you that? Uncovering the source of believed cues to deception. *International Journal of Psychological Studies, 6,* 19–32. https://doi.org/10.5539/ijps.v6n1p19.

Hurst, M., & Oswald, M. (2012). Mechanisms underlying response bias in deception detection. *Psychology, Crime & Law, 18,* 759–778. https://doi.org/10.1080/1068316X.2010.550615.

Jussim, L. (2012). *Social perception and social reality: Why accuracy dominates bias and self-fulfilling prophecy.* Oxford, UK: Oxford University Press.

King, W. R., & Dunn, T, M. (2010). Detecting deception in field settings. A review and critique of the criminal justice and psychological literatures. *Policing: An International Journal of Police Strategies & Management, 33,* 305–320. https://doi.org/10.1108/13639511011044902.

Kuhn, T. S. (1962a). Historical structure of scientific discovery. *Science, 136*(3518), 760–764.

Kuhn, T. S. (1962b). *The structure of scientific revolutions.* Chicago, IL: University of Chicago Press.

Lee, K., Cameron, C. A., Doucette, J., & Talwar, V. (2002). Phantoms and fabrications: Young children's detection of implausible lies. *Child Development, 73,* 1688–1702. https://doi.org/10.1111/1467-8624.t01-1-00499.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125,* 255–275. https://doi.org/10.1037/0033-2909.125.2.255.

Levine, T. R. (2010). A few transparent liars: Explaining 54% accuracy in deception detection experiments. In C. Salmon (Ed.), *Communication Yearbook 34* (pp. 40–61). Thousand Oaks, CA: Sage. https://doi.org/10.1080/23808985.2010.11679095.

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392. https://doi.org/10.1177/0261927X14535916.

Levine, T. R., Asada, K. J. K., & Park, H. S. (2006). The lying chicken and the gaze avoidant egg: Eye contact, deception, and causal order. *Southern Communication Journal, 71,* 401–411. https://doi.org/10.1080/10417940601000576.

Lilienfeld, S. O., Lynn, S. J., Ruscio, J., & Beyerstein, B. L. (2010). *50 great myths of popular psychology: Shattering widespread misconceptions about human behavior.* Chichester, UK: Wiley.

Lloyd, E. P., Deska, J. C., Hugenberg, K., McConnell, A. R., Humphrey, B. T., & Kunstman, J. W. (2019). Miami University deception detection database. *Behavior Research Methods, 51*(1), 429–439. https://doi.org/10.3758/s13428-018-1061-4.

Lord Laming. (2009). *The protection of children in England: A progress report.* London, UK: The Stationery Office.

Mandelbaum, E. (2014). Thinking is believing. *Inquiry: An Interdisciplinary Journal of Philosophy, 57*(1), 55–96. https://doi.org/10.1080/0020174X.2014.858417.

Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology, 89,* 137–149. https://doi.org/10.1037/0021-9010.89.1.137.

Masip, J. (2014). Generalized Communicative Suspicion. In T. Levine (Ed.), *Encyclopedia of deception* (Vol. 1, pp. 404–406). Thousand Oaks, CA: Sage.

Masip, J., Alonso, H., Garrido, E., & Antón, C. (2005). Generalized Communicative Suspicion (GCS) among police officers: Accounting for the investigator bias effect. *Journal of Applied Social Psychology, 35,* 1046–1066. https://doi.org/10.1111/j.1559-1816.2005.tb02159.x.

Masip, J., Alonso, H., Garrido, E., & Herrero, C. (2009). Training to detect what? The biasing effects of training on veracity judgments. *Applied Cognitive Psychology, 23,* 1282–1296. https://doi.org/10.1002/acp.1535.

Masip, J., Alonso, H., Herrero, C., & Garrido, E. (2016). Experienced and novice officers' Generalized Communication Suspicion and veracity judgments. *Law and Human Behavior, 40,* 169–181. https://doi.org/10.1037/lhb0000169.

Masip, J., & Garrido, E. (2001, June). *Experienced and novice officers' beliefs about indicators of deception.* Paper Presented at the 11th European Conference of Psychology and Law, Lisbon, Portugal.

Masip, J., Garrido, E., Herrero, C., Antón, C., & Alonso, H. (2006). Officers as lie detectors. Guilty before charged. In D. Chadee & J. Young (Eds.), *Current themes in social psychology* (pp. 187–205). Mona, Jamaica: The University of the West Indies Press.

Masip, J., & Herrero, C. (2017). Examining police officers' response bias in judging veracity. *Psicothema, 29,* 490–495. https://doi.org/10.7334/psicothema2016.357.

Mayo, R., Schul, Y., & Burnstein, E. (2004). "I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology, 40,* 433–449. https://doi.org/10.1016/j.jesp.2003.07.008.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26,* 469–480. https://doi.org/10.1023/A:1020278620751.

Nadarevic, L., & Erdfelder, E. (2013). Spinoza's error: Memory for truth and falsity. *Memory and Cognition, 41,* 176–186. https://doi.org/10.3758/s13421-012-0251-z.

Nahari, G. (2012). Elaborations on credibility judgments by professional lie detectors and laypersons: Strategies of judgment and justification. *Psychology, Crime & Law, 18,* 567–577. https://doi.org/10.1080/1068316X.2010.511222.

Nahari, G. (2017). Top-down processes in interpersonal reality monitoring assessments. *Psychology, Public Policy, and Law, 23,* 232–242. https://doi.org/10.1037/law0000110.

Nahari, G., Vrij, A., & Fischer, R. P. (2014). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology, 19,* 227–239. https://doi.org/10.1111/j.2044-8333.2012.02069.x.

O'Brien, B. (2009). Prime suspect: An examination of factors that aggravate and counteract confirmation bias in criminal investigations. *Psychology, Public Policy, and Law, 15,* 315–334. https://doi.org/10.1037/a0017881.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), 943–951. https://doi.org/10.1126/science.aac4716.

Paoline, E. A., III. (2003). Taking stock: Toward a richer understanding of police culture. *Journal of Criminal Justice, 31,* 199–214. https://doi.org/10.1016/s0047-2352(03)00002-3.

Platzer, C., & Bröder, A. (2012). Most people do not ignore salient invalid cues in memory-based decisions. *Psychonomic Bulletin & Review, 19,* 654–661. https://doi.org/10.3758/s13423-012-0248-4.

Pohl, R. F. (2011). On the use of recognition in inferential decision making: An overview of the debate. *Judgment and Decision Making, 6,* 423–438.

Porter, S., Woodworth, M., & Birt, A. R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and Human Behavior, 24,* 643–658. https://doi.org/10.1023/A:1005500219657.

Reyes-Jaquez, B., & Echols, C. H. (2015). Playing by the rules: Self-interest information influences children's trust and trustworthiness in the absence of feedback. *Cognition, 134,* 140–154. https://doi.org/10.1016/j.cognition.2014.10.002.

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology, 96,* 538–588. https://doi.org/10.1037/a0014038.

Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbook of personality and social psychology: Attitudes and social cognition* (pp. 203–229). Washington, DC: APA.

Schwarz, N., Newman, E., & Leach, W. (2016). Making the truth stick and the myths fade: Lessons from cognitive psychology. *Behavioral Science & Policy, 2,* 85–95. https://doi.org/10.1353/bsp.2016.0009.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*, 367–447. https://doi.org/10.1017/s0140525x00035032.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41,* 1–19. https://doi.org/10.1146/annurev.ps.41.020190.000245.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford, UK: Blackwell.

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology, 7,* 420. https://doi.org/10.3389/fpsyg.2016.00420.

Sporer, S. L., Masip, J., & Cramer, M. (2014). Guidance to detect deception with the Aberdeen Report Judgment Scales: Are verbal content cues useful to detect false accusations? *American Journal of Psychology, 127,* 43–61. https://doi.org/10.5406/amerjpsyc.127.1.0043.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20,* 421–446. https://doi.org/10.1002/acp.1190.

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13,* 1–34. https://doi.org/10.1037/1076-8971.13.1.1.

Street, C. N. H. (2015). ALIED: Humans as adaptive lie detectors. *Journal of Applied Research in Memory and Cognition, 4,* 335–343. https://doi.org/10.1016/j.jarmac.2015.06.002.

Street, C. N. H., Bischof, W. F., Vadillo, M. A., & Kingstone, A. (2015). Inferring others' hidden thoughts: Smart guesses in a low diagnostic world. *Journal of Behavioral Decision Making, 29,* 539–549. https://doi.org/10.1002/bdm.1904.

Street, C. N. H., & Kingstone, A. (2016). Aligning Spinoza with Descartes: An informed Cartesian account of the truth bias. *British Journal of Psychology, 108*(3), 453–466. https://doi.org/10.1111/bjop.12210.

Street, C. N. H., & Richardson, D. C. (2015a). Descartes versus Spinoza: Truth, uncertainty, and bias. *Social Cognition, 33,* 227–239. https://doi.org/10.1521/soco.2015.33.2.2.

Street, C. N. H., & Richardson, D. C. (2015b). Lies, damn lies, and expectations: How base rates inform lie-truth judgments. *Applied Cognitive Psychology, 29,* 149–155. https://doi.org/10.1002/acp.3085.

Street, C. N. H., & Richardson, D. C. (2015c). The focal account: Indirect lie detection need not access unconscious, implicit knowledge. *Journal of Experimental Psychology: Applied, 21,* 342–355. https://doi.org/10.1037/xap0000058.

Street, C. N. H., & Vadillo, M. A. (2016). Can the unconscious boost lie-detection accuracy? *Current Directions in Psychological Science, 25,* 246–250. https://doi.org/10.1177/0963721416656348.

Strömwall, L., Granhag, P. A., & Hartwig, M. (2004). Practitioners' beliefs about deception. In P.-A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 229–250). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511490071.010.

Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology, 45,* 74–83. https://doi.org/10.1037/0022-3514.45.1.74.

Vrij, A. (2004). Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology, 9,* 159–181. https://doi.org/10.1348/1355325041719356.

Vrij, A. (2008). *Detecting lies and deceit. Pitfalls and opportunities.* Chichester, UK: Wiley.

CHAPTER 13

A Model of Trust Manipulation: Exploiting Communication Mechanisms and Authenticity Cues to Deceive

Emma J. Williams and Kate Muir

A crucial factor when successfully deceiving others is to instill trust in the receiver that the professed scenario and communication is genuine. In order to achieve this, deceivers often manipulate established norms and trust mechanisms to aid in their deception. This chapter explores the role of trust in deceptive communications and the different ways in which trust may be manipulated by deceivers to influence the judgments of message recipients. Following this analysis, we present an initial model of trust manipulation that brings all of these factors together to consider how elements of communication, such as building rapport and the use of authenticity cues, may be used to invoke trust in order to effectively deceive others.

E. J. Williams (✉)
School of Experimental Psychology, University of Bristol, Bristol, UK
e-mail: emma.williams@bristol.ac.uk

K. Muir
School of Management, University of Bath, Bath, UK
e-mail: K.Muir@bath.ac.uk

Introducing Trust

What Is Trust?

Trust is a term most people use frequently in everyday conversation. We talk about trusting individuals, or trusting a particular brand or organization more than another. We can even talk about trust in relation to abstract concepts, such as trusting in love or in fate. But what do we mean when we talk about trust? It is a topic studied across many disciplines, yet there is no single definition of trust.

For psychologists, trust is a state of the individual: a psychological state comprising the intention to accept vulnerability, based upon positive expectations of the intentions and behavior of another (Rousseau, Sitkin, Burt, & Camerer, 1998). Trust is also defined as an expression of confidence in another person, that one will not be put at risk or harmed by their actions. Trust develops over time as a result of experiences that show a person's behavior is predictable and dependable (Kerr, Stattin, & Trost, 1999). In contrast, for sociologists trust is defined by the quality of the relationship between individuals. Predictability is one way to quantify that relationship, in terms of the probability with which an actor assesses that another actor will behave in a certain way (Luhmann, 1979). If our predictions or expectations about the behavior of another come true, this leads to trust in the individual. Other definitions describe trust as including "goodwill": mutual expectations of reciprocity between partners which leads to cooperation and trust (Ring & Van de Ven, 1992).

Scales developed to objectively measure trust also define it in slightly different ways. Rotter's (1967) interpersonal trust scale defined trust as the "*expectancy held by an individual or group that the word, promise, verbal or written statement of another individual or group can be relied upon*". This scale measures the general tendency to trust other people and society, but not individual people. A scale developed to measure trust in specific people (e.g., friend or partner) developed by Rempel, Holmes, and Zanna (1985) defines trust in terms of predictability, dependability, and faith.

Regardless of discipline or measurement, definitions of trust seemingly center around positive expectations of behavior being met by another. Given it is impossible for us to always know everything about someone else (after all, we cannot read minds), trust is vital to the functioning of individual relationships, organizations, and social systems. In this chapter, we will explore some of the methods used by individuals to develop trust in communication, and signal that the positive expectations held by another will be met. We then discuss how those intending to deceive can exploit these communicative mechanisms.

How Is Trust Communicated?

Individuals and organizations can communicate their trustworthiness by the use of *symbols* and *signals*. Communicative *symbols* are indicators of trust that enable quick evaluations without expending major cognitive resources. One example of communicative symbols of trust is the uniforms worn by professionals. Uniforms, such as the white coats worn by doctors, or the uniforms worn by police, act as a shortcut to the characteristics of the individual wearing it. A uniform signifies respect, authority, and knowledge in a particular field. Most individuals in a society recognize a uniform as representing membership of a profession, and what the profession stands for. It implicitly signifies reliability; we expect that an individual wearing a medical uniform, for instance, will act to protect our health, whether by providing information and advice or administering medical aid. Thus, communicative symbols operate as signs of trust on the basis of shared understanding of what the symbol represents. We form positive expectations of behavior based on the symbol alone; we trust what the symbol represents, without necessarily knowing anything else about the individual or organization.

Communicative symbols can be of particular importance in the online environment where other communicative cues are limited. Online, symbols of trust can take the form of ratings or rankings (e.g., stars), or a well-known brand name or logo on a vendor's website. The image of a lock at the bottom of a webpage, indicating a secure transaction, can act as a symbol of trust in the vendor's webpage, as can a professional webpage appearance (Fung & Lee, 1999). Symbols of trust therefore act as heuristics, enabling a quick and easy evaluation of trust in both face-to-face and online communicative environments.

Communicative *signals*, on the other hand, form part of our socially shared signaling system in which information is encoded into signals, which are sent by one individual and received and decoded by the receiver. Smiling is one example; a smile is a social signal sent out by an individual and decoded by the receiver as representing happiness, joy, pleasure, congratulations, and so on. Communicative signals of trust, therefore, can be sent out by one individual and detected and decoded by another. Such signals can take various forms. Gaze, for instance, can act as a signal of trust. Gaze-cuing studies, whereby participants are presented with several faces looking toward or away from a target, suggest that people who make direct eye contact are judged as more likable and trustworthy than those who do not make eye contact (Bayliss & Tipper, 2006). Similarly, a study conducted by Mason, Tatkow, and Macrae (2005) found that faces looking directly toward participants were judged as more physically attractive and likable than faces looking away from them. Faces which consistently look in the direction of a target object (known as predictive gaze) have also been shown to be judged as more

trustworthy than faces in which gaze is not predictive (Bayliss & Tipper, 2006). Individuals who make direct eye contact could, therefore, be (intentionally or not) sending a signal of trustworthiness. The form and characteristics of the message itself can also act as a signal of trust. People trust the sender of a message more when the language of the message is structured and elaborated compared to unelaborated (Lev-Ari & Keysar, 2010), and includes complicated language (Thiebach, Mayweg-Paus, & Jucks, 2015). These message characteristics indicate expertise and can thus act as a communicative signal of trustworthiness ("trust me, I know what I am doing").

People rely on communicative symbols and signals to form judgments of trust. This can particularly be the case in the early stages of a relationship, where there may be a lack of other evidence regarding the trustworthiness of a partner based on past experience. Crucially, communicative signals and symbols can be deliberately exploited to convey a façade of trust, without any intention to reciprocate or meet positive expectations of behavior.

Developing Trust in Communication

Language Mechanisms to Develop Trust

Whether we intend it to or not, the language we use communicates a great deal about our personality, attitudes, and intentions—including our trustworthiness (e.g., Pennebaker, 2011; Tausczik & Pennebaker, 2010). How we use language can act as both a powerful signal of trust and a mechanism to foster trust, rapport, and liking in communication. Such linguistic mechanisms can be used strategically by those intending to foster trust to deceive.

Self-disclosure is one example of a linguistic mechanism or communication strategy which supports trust-building processes. Self-disclosure is defined as the "process of making the self known to others" (Jourard & Lasakow, 1958, p. 91) or the voluntary presentation of information that is of an intimate or personal nature to another person (Miller, 2002) and has been demonstrated to foster interpersonal trust. Ensari and Miller (2002) found that self-disclosure by a member of a perceived out-group induced feelings of trust toward the out-group. Similarly, self-disclosure has been associated with increased trust between in- and out-group members which in turn promoted positive attitudes toward the out-group (Turner, Hewstone, & Voci, 2007). Self-disclosure also helps to foster rapport and trust during negotiations. Exchanging personal information during a short informal "schmoozing" session prior to a negotiation has been shown to lead to more rapport and trust before and after the negotiation, and more positive outcomes of the negotiation for both parties (Morris, Nadler, Kurtzberg, & Thompson, 2002).

One theory about how self-disclosure works to engender trust is that providing personal information (i.e., self-disclosure) demonstrates a willingness to be vulnerable, and can be understood as an offer of trust which is often

reciprocated (Jourard, 1971). Further, receiving personal information implies being trusted by the discloser, and people like and trust those who trust them (Petty & Mirels, 1981). When people disclose to us, we disclose more in return, leading to mutual interpersonal attraction (Worthy, Gary, & Kahn, 1969). Thus, self-disclosure is part of the process of developing trust in relationships. As people get to know one another, they realize the information that they disclose is safe and so trust in each other increases. Deceivers can use this aspect of communication to deceive; by self-disclosing and encouraging self-disclosure in others, they can foster feelings of trust.

Self-disclosure, and manipulating situations which encourage self-disclosure, may be considered just one tool in a box of linguistic tricks that deceivers can employ to engender trust. Verbal mimicry is another such linguistic trick. Verbal mimicry describes the mimicry of one's conversational partner's speech, such as imitating accents or speech rate (Giles & Powesland, 1975; Webb, 1969). Verbal mimicry is associated with positive outcomes of conversations and can be used strategically to enhance interpersonal connections and relationships. For example, verbal mimicry by sales assistants has been associated with bigger tips (van Baaren, Holland, Steenaert, & van Knippenberg, 2003), higher sales, and more positive perceptions of the mimicker (Jacob, Gueguen, Martin, & Boulbry, 2011). Mimicry of verbal expressions by participants in a speed-dating scenario has been associated with higher evaluations of attractiveness of the mimicker (Guéguen, 2009). Verbal mimicry is thus positively associated with both quantifiable and subjective results for the mimicker.

These positive effects of verbal mimicry extend to trust also. Within text-based instant messaging sessions, pairs who mimicked each other's language were more likely to trust each other (Scissors, Gill, & Gergle, 2008; Scissors, Gill, Geraghty, & Gergle, 2009). Importantly, research shows that verbal mimicry can be leveraged strategically to increase trust. Individuals negotiating using instant messaging (text-based chat) who were instructed to intentionally mimic the language of their negotiation partner found that this increased levels of trust, which in turn predicted more positive outcomes of the negotiation (Swaab, Maddux, & Sinaceur, 2011).

Verbal mimicry is part of the wider phenomenon of behavioral mimicry. People naturally mimic the facial expressions, postures, and behaviors of their conversational partners (for a review, see Chartrand, Maddux, & Lakin, 2006). Like verbal mimicry, behavioral mimicry is associated with increased liking (Chartrand & Bargh, 1999) and can be intentionally performed to foster feelings of trust. Maddux, Mullen, and Galinsky (2008) instructed half the participants in a negotiation exercise to mimic the mannerisms of their negotiating partner. Participants who mimicked gained a better deal at the end of the negotiation, and this effect was mediated by increased feelings of trust experienced by the mimicked partner.

How does mimicry work to increase trust? Mimicry is thought to be non-conscious, in that people do it without realizing they are mimicking or

being mimicked. So, mimicry is thought to serve as a sort of "social glue", bonding people together, helping them to affiliate, feel connected, and form rapport (Chartrand & Dalton, 2009). In support of this, being mimicked has led to people defining themselves in terms of their relationships with others (Stel & Harinck, 2011). Mimicry also acts to communicate understanding, fostering an empathetic bond between the mimicker and the mimicked. Stel and Vonk (2010) found that when one individual mimics another, both parties become more emotionally attuned, in that they feel similar emotions. It is likely that this contributes to feelings of trust. Indeed, some research suggests that correctly inferring the feelings of others and giving appropriately supportive responses increases trust in online situations (Feng, Lazar, & Preece, 2004). Thus, mimicry, both verbal and behavioral, can act in several ways as a signal of trust. It signals that "I am on your wavelength, and I understand you – I am trustworthy". In this way, verbal and behavioral mimicry can be intentionally manipulated to enhance rapport and feelings of affiliation, and foster a façade of trust.

Situational Mechanisms to Develop Trust

The particular linguistic aspects of a communication are unlikely to be the sole determinants of evaluations of trustworthiness, however. People's expectations of the scenario and the context in which they are operating can also impact how incoming information is considered and evaluated (Williams, Beardmore, & Joinson, 2017). One mechanism by which such situational factors may impact evaluations is by influencing the depth of cognitive processing that an individual engages in when a communication is encountered. Both the heuristic-systematic model (Eagly & Chaiken, 1993) and the elaboration likelihood model (Petty & Cacioppo, 1986) suggest that the relative persuasiveness of a particular message is dependent on the processing strategy that is used by the message recipient, namely whether they rely on relatively automatic, heuristic forms of processing or more resource-intensive, systematic processing strategies. The use of heuristic-based strategies has been linked to a number of automatic biases in decision-making, including basing judgments on pre-conceived stereotypes, expectations, or emotional responses (Kahneman, 2011; Tversky & Kahneman, 1974).

The potential role of heuristics in relation to evaluations of trust or suspicion has been explicitly highlighted in the deception field through the proposed existence of a truth bias within the general population (Bond & DePaulo, 2006; Levine, 2014). This suggests that, in general, people assume that information that they encounter in the environment is more likely to be true than false, unless there is a particular reason to doubt its legitimacy. For instance, in scenarios where deception is considered to be more likely, such as sales contexts or police interviews, this truth bias has been found to diminish (DePaulo & DePaulo, 1989; Meissner & Kassin, 2002). Similarly, basing

considerations of likely legitimacy on pre-conceived stereotypes regarding what a deceptive communication typically "looks like" can also be related to these heuristic forms of processing. In this respect, whether a message sender appears honest, or at least matches our expectations of what we consider honest people to look like, can impact evaluations of trustworthiness, regardless of whether the individual is actually telling the truth (Bond & DePaulo, 2008).

Research has also identified a number of social influence processes that can impact how people respond to communicative scenarios across different contexts (Cialdini, 2007). Again, these represent relatively automatic and ingrained decision biases, including *compliance with authority* (i.e., people's tendency to comply with requests from authority figures), *social conformity* (i.e., people's tendency to conform to the behavior of those around them), *consistency and commitment* (i.e., people's desire to remain consistent with their previous behavior and commitments), and being more easily influenced by those that we *like* and those that we consider ourselves to be *similar* to. In this way, the interaction between a message receiver and a message sender, including evaluations of relative trustworthiness, can be situated within a wider social context. For instance, message senders who are perceived as authority figures within a particular scenario may be considered as more likely to be legitimate, and therefore accurate, sources of information, leading to the authenticity of message content being assumed. Similarly, information that comes from perceived similar others, or that is broadly considered to be genuine by relevant peers, may be evaluated as likely to be trustworthy. People are known to expect others to be similar to themselves in terms of their beliefs, characteristics, and traits (a tendency termed *self-projection*; Robbins & Krueger, 2005), and therefore, this perception of similarity can easily be exploited by a message sender in order to increase their relative influence over another individual. Indeed, identifying with the character and situation of potential protagonists within any professed scenario is considered important if emotional responses, such as empathy, are to be evoked in the message recipient (Hoeken & Sinkeldam, 2014).

The relative impact of social influence processes on response behavior has also been shown to be dependent on wider contextual factors related to the particular scenario encountered. For instance, individuals in established positions of social power have been found to be more self-focused, more resistant to social influence, and more likely to act in accordance with their personal preferences (Pitesa & Thau, 2013). It has been suggested that this occurs because such individuals are less reliant on other people for maintaining their relative position within a company, making them less dependent on others and thus allowing them to focus their attention internally. Individuals in less powerful positions, however, are more dependent on external factors and circumstances for their success and are, therefore, more likely to focus on contextual stimuli, such as social influence, in order to increase the predictability

of their surrounding environment (Guinote, 2007). Although such propositions are based on relatively stable, situational factors (i.e., one's relative position within an organization), more transient notions of power may also have an effect, with message receivers potentially operating in contexts where they are temporarily in a position of low social power compared to the professed message sender (Lee & Soberon-Ferrer, 1997).

Since the majority of people simply do not have the mental resource and time required to systematically consider the authenticity of all communications that they encounter, it is perhaps unsurprising that we must instead rely on quick decision rules and biases to help us to identify when more in-depth evaluation of a message is required. Unfortunately, it is precisely these decision rules that deceivers attempt to manipulate in order to increase the likelihood that their professed scenario will be perceived as genuine. In the following section, the various mechanisms by which such trust manipulation can occur are considered, focusing on how these concepts are differentially applied in a range of fraudulent communications.

MANIPULATING TRUST TO DECEIVE

Action Fraud, the UK's national fraud and cybercrime reporting center, defines fraud as "trickery used to gain a dishonest advantage, which is usually financial, over another person" (Action Fraud, 2017). This definition encompasses what are more commonly known as scams, whereby some form of deception is used to persuade a message receiver that a fraudulent communication can be trusted and that the professed scenario is legitimate and should be responded to. By invoking trust in the legitimacy of the communication, whether it be online, via email, through the post, over the telephone, or face-to-face, message recipients are more likely to be persuaded to respond (i.e., to click on malicious links, make a monetary payment, or provide personal information), usually in order to achieve a potential gain (e.g., monetary prizes, romantic relationships, desirable opportunities) or to avoid some form of loss, such as securing an account or maintaining computer access (Atkins & Huang, 2013; Workman, 2008).

In order to achieve this outcome, a number of the mechanisms discussed in the previous section are often incorporated within fraudulent communications. These primarily rely on encouraging message receivers to engage in heuristic processing strategies rather than considering the content of the communication in-depth, thus increasing the chances that people will default to a truth bias and assume that the message is genuine (Levine, 2014). The use of logos and other design cues allows visual communications, such as phishing emails and websites, to effectively mimic recognizable brands and reputable organizations. These visual and design aspects have been shown to influence legitimacy judgments, with initial decisions based on these 'heuristic' elements at the expense of more systematic consideration of website content

and authenticity cues, such as the URL (Dhamija, Tygar, & Hearst, 2006; Fogg et al., 2002; Sillence, Briggs, Fishwick, & Harris, 2006). Similar cues may also be used in offline scenarios, whereby fraudsters may wear particular clothing (e.g., wearing a high-visibility jacket) or use props, such as ID badges, that are associated with authentic scenarios. In this way, wider aspects of the scenario outside of the linguistic content of the message itself can be designed in order to match people's expectations regarding legitimate information, thus reducing the likelihood that suspicion will be triggered.

Fraudulent communications also include a range of influence techniques to encourage people to respond to messages quickly and without thinking, and thus enhance the likelihood that heuristic processing strategies will be used (Stajano & Wilson, 2011). For instance, communications often attempt to invoke a sense of urgency in message recipients by including explicit deadlines or creating scenarios that evoke a degree of time pressure. They may mimic organizations or individuals that are considered to have a degree of authority, such as government institutions, law enforcement agencies, or senior managers within organizations. Crucially, they also often invoke some form of emotion in the message receiver, which can be either positive (e.g., excitement at a potential opportunity, desire for a romantic relationship, hope for a potential miracle cure, or curiosity to know more) or negative (e.g., anxiety about a potential security breach on your account, guilt regarding not helping another in need, or fear of the potential negative ramifications of not responding).

The use of influence techniques within fraudulent communications has been suggested to impact the likelihood that message receivers will notice suspicious elements within a communication, with Vishwanath, Herath, Chen, Wang, and Rao (2011) suggesting that urgency cues within phishing emails can monopolize limited attentional resources and lead to authenticity cues, such as an inaccurate sender address, being overlooked. More recently, Vishwanath, Harrison, and Ng (2016) have proposed that the use of heuristic processing strategies when viewing phishing emails decreases the likelihood that suspicious elements within the email will be noticed. If such elements are neglected, then suspicion with regard to the legitimacy of the message will not be triggered, leading to message receivers assuming that a communication is genuine (Levine, 2014). Unless individuals already have a reason to be suspicious, potentially as a result of enhanced awareness of risks and/or the relative likelihood of deception, then they are considered to be more likely to rely on heuristic processing strategies when making their decisions since more resource-intensive, systematic strategies are not considered necessary or worthwhile (Boureau, Sokol-Hessner, & Daw, 2015; Vishwanath et al., 2016).

Whereas communications such as phishing emails effectively "hijack" existing trust in a particular organization or individual, other types of deceptive communications rely on building trust from scratch. For example, online

romance scammers must create relationships with others using online dating websites; rather than mimicking an existing relationship, they must create a new one. To do this, they exploit a range of stereotypes in creating profiles that are likely to be considered as attractive and likable by other users. Typically, these include using fake profile photographs that are physically attractive and creating narratives that are likely to be considered desirable by potential partners, such as older, widowed businessmen or military personnel, and younger women (Whitty, 2015). These profiles also often attempt to evoke a degree of similarity with potential victims.

Typically, online romance scams create intense relationships with their victims, engaging in regular communication, using early protestations of love and moving their communications off of the original dating website to decrease the likelihood that their activities will be monitored. Plausible scenarios may be created to avoid a face-to-face meeting, such as working or living temporarily overseas or serving with the military, while attempts at video calls may be thwarted by apparently poor Internet connections. Both social projection processes (Robbins & Krueger, 2005) and the online disinhibition effect (Suler, 2004), whereby people feel able to self-disclose more personal information in online settings than they typically would in face-to-face contexts, likely aid in the development of trust within such relationships. Once a relationship is established, a crisis is then created that requires funds from the victim, such as an emergency hospital visit while abroad, or lost documentation. These scenarios are often designed to be emotionally evocative, such that any doubts that may occur regarding the legitimacy of the scenario will be overcome by a desire to help, guilt regarding potentially refusing the request, and social norms in relation to supporting relational others, helping those in need and reciprocity.

The various forms of deception within fraudulent communications, in particular the various mediums that they can use, result in substantial variation in the potential authenticity cues that may be available (or not available) to determine whether the scenario is legitimate, and the factors that may impact whether this authenticity information is even used. Previous explanations of the generally poor ability of most people to detect deception have considered this in relation to a reliance on default heuristics and biases when making decisions and, more recently, on the role of the wider context in impacting the extent to which these biases may occur (Street, 2015). The recently proposed adaptive lie detector (ALIED) (Street, 2015) theory suggests that the low availability of cues to detect the majority of deceptive communications within the environment is a major factor influencing people's judgments. This lack of definitive cues to deception means that people must instead base their decisions on expectations and previous experience regarding the likelihood that they will encounter deception in any given context. Since the majority of people are considered to tell the truth the majority of the time, this leads to an overall truth bias in judgments. Aspects of the particular context are then used to determine whether deception is more likely to be encountered

in a particular scenario, and if this is the case, the incidence of truth bias decreases.

In accordance with the ALIED theory above, when authenticity cues are not present within a communication then people will be more likely to rely on aspects of the wider context to inform their judgments. In this way, individual beliefs regarding cyber risks, the degree of knowledge that people have, and their perceptions of personal susceptibility to deceptive communications online, may all influence how much people trust communications that they encounter (Vishwanath et al., 2016). However, the impact of context could also impact judgments of trustworthiness in other ways. For instance, situations where the availability of cognitive resource to systematically consider information is reduced may result in people missing potential authenticity cues, even when they are available (Williams, Morgan, & Joinson, 2017). Similarly, the existence of strong habits in relation to online communications (e.g., responding relatively automatically to emails and being less able to control this behavior) has also been shown to reduce the likelihood that suspicious elements will be noticed or considered (Vishwanath, 2015).

An Initial Model of Trust Manipulation

The previous sections of this chapter highlight a number of language mechanisms and situational mechanisms that may influence trust judgments, providing examples of how these mechanisms may be manipulated by nefarious individuals when attempting to deceive others. Within this section, these aspects are combined and presented within an initial model of trust manipulation (see Fig. 13.1), providing a foundation for future experimental work exploring the potential interrelationships of these factors. This focuses on
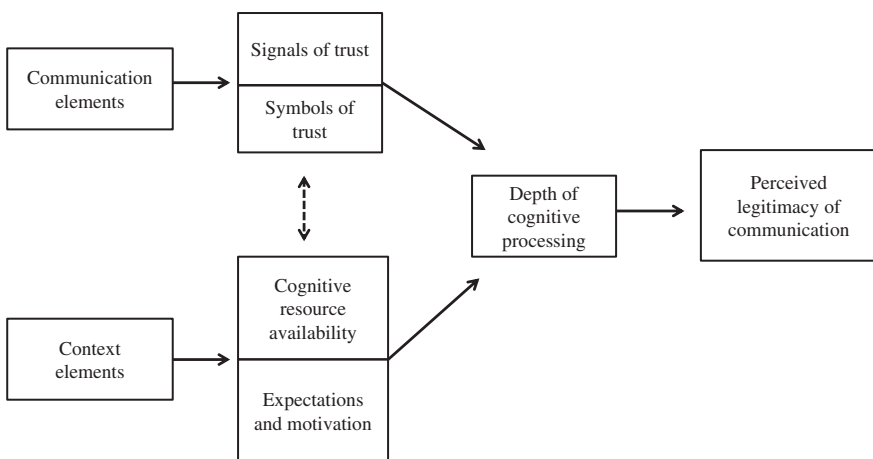
Fig. 13.1   An initial model of how receiver trust may be manipulated by deceivers

aspects of both the *communication* itself and the wider *context* in which the communication is received, particularly how these aspects may influence the *depth of cognitive processing* that is engaged in when considering deceptive communications. Whether these communications contain identifiable cues of the deception or do not contain any such cues, the processing strategies that are used when the information is encountered are considered fundamental in influencing the overall *perceived legitimacy* of the communication.

When creating a deceptive communication, a range of communicative *symbols* and *signals* are often invoked within the *communication* that traditionally signals trust. These may range from aspects of the design of a fraudulent email (such as the use of recognizable logos or branding linked to trusted institutions), to the use of particular words or body language that is considered suggestive of trustworthy communications, individuals, or scenarios. These symbols and signals are interpreted by the receiver in the context of the particular scenario, with different aspects of the communication likely to be more or less systematically evaluated according to the degree to which deception is considered likely to be encountered. For instance, when individuals are expecting a particular email from a particular sender, then fraudulent emails that effectively mimic these expectations may be less likely to be noticed, even when legitimacy cues are present, due to insufficient suspicion in the potential for deception to occur. Effectively, people assume that the communication is genuine due to this matching of expectations. As a result, signals and symbols may be differentially interpreted, and also differentially evaluated, when making response decisions across different scenarios.

The use of these symbols and signals in a deceptive communication, therefore, must be considered in relation to *elements of the wider context*. Where message receivers are likely to be more suspicious of a communication, and therefore more *motivated* to engage in resource-intensive, systematic processing, symbols and signals may be considered more important in the creation of the communication in order to develop and retain trust. In this case, the use of such trust cues may either (a) encourage more heuristic-level processing when the communication is encountered (e.g., the presence of particular wording or branding), and thus increase the likelihood that the communication will still be considered genuine, or (b) increase the likelihood that the communication will still be judged as genuine even after more systematic processing has been engaged in. If the latter is correct, then such cues effectively serve to increase the resistance of the deceptive communication to being identified when exposed to more in-depth investigation.

However, in contexts where individuals are more likely to assume that the communication is genuine, the emphasis on symbols and signals by message senders may be reduced, with the requirement that these aspects are merely sufficient to prevent suspicion being aroused in the first place. As mentioned previously, such assumptions of legitimacy may occur in environments where numerous similar types of genuine communications are encountered, thus

exploiting elements of *expectations* or message familiarity. In this scenario, individuals may not consider systematic evaluation of the legitimacy of such messages to be an effective use of limited cognitive resources, reducing their *motivation* to use such processing strategies. Alternatively, assumptions of authenticity may be relied upon as a result of limitations within the wider environment, rather than due to internal motivations per se. For example, when people are busy or are under some form of cognitive pressure, then the *availability of cognitive resource* to engage in more systematic processing may be reduced, thus increasing the likelihood that heuristic forms of processing will be relied upon when making decisions.

Conclusions and Future Directions

Within this chapter, a number of factors have been identified that may impact the extent to which individuals trust a particular communication that they encounter. These factors include aspects related to the communication itself as well as the sender of that communication, which can be combined with characteristics of the wider context in which the communication is received to consider how judgments of trust may be influenced by deceivers. For instance, the extent to which people use authenticity cues when making decisions regarding message legitimacy is likely to be impacted by a range of factors related not only to the specific communication itself, but also to aspects of the wider environment in which the communication is encountered. An initial model of trust manipulation has been presented, which combines aspects of a communication that may be used to build rapport or symbolize authenticity and factors related to wider situational aspects that may impact how people respond to a communication, such as the use of particular social influence cues, matching current expectations and norms within the environment, and the creation of pressured scenarios that are likely to be particularly relevant to individuals in certain contexts. These factors are considered to impact evaluations via influencing the depth of processing that an individual engages in when considering how to respond to information (Levine, 2014; Street, 2015; Vishwanath et al., 2016).

Researchers increasingly recognize the importance of identifying potential relationships between contextual, individual, and message-related factors when considering how people make judgments regarding the legitimacy of communications. Understanding and investigating this complexity represents a considerable challenge. However, it is vital if effective interventions are to be developed that are able to reduce individual susceptibility to deceptive communications across a variety of contexts, both when cues to deception may be lacking and when they are present but underused by message recipients. It is hoped that this model will provide an initial basis for further discussion and investigation of the variety of factors that may impact people's authenticity decisions. In particular, the precarious balance between trust

and suspicion of communications, and the relative thresholds that people use to determine whether a communication is more or less likely to be genuine across various contexts.

References

Action Fraud. (2017). What is fraud and cyber crime. Cited at http://www.action-fraud.police.uk/what-is-fraud, on February 28, 2017.

Atkins, B., & Huang, W. (2013). A study of social engineering in online frauds. *Open Journal of Social Sciences, 1*(3), 23–32. https://doi.org/10.4236/jss.2013.13004.

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgements: Should eye trust you? *Psychological Science, 17*(6), 514–520. https://doi.org/10.1111/j.1467-9280.2006.01737.x.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477–492. https://doi.org/10.1037/0033-2909.134.4.477.

Boureau, Y. L., Sokol-Hessner, P., & Daw, N. D. (2015). Deciding how to decide: Self-control and meta-decision making. *Trends in Cognitive Science, 19*(11), 700–710. https://doi.org/10.1016/j.tics.2015.08.013.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893.

Chartrand, T. L., & Dalton, A. N. (2009). Mimicry: Its ubiquity, importance, and functionality. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 458–483). New York, NY: Oxford University Press.

Chartrand, T. L., Maddux, W. W., & Lakin, J. L. (2006). Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. Uleman, & J. A. Bargh (Eds.), *The new unconscious.* New York: Oxford University Press.

Cialdini, R. (2007). *Influence: The psychology of persuasion.* New York: HarperCollins.

DePaulo, P. J., & DePaulo, B. M. (1989). Can attempted deception by salespersons and customers be detected through nonverbal behavioral cues? *Journal of Applied Social Psychology, 19,* 1552–1577. https://doi.org/10.1111/j.1559-1816.1989.tb01463.x.

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 581–590). https://doi.org/10.1145/1124772.1124861.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt.

Ensari, N., & Miller, N. (2002). The out-group must not be so bad after all: The effects of disclosure, typicality, and salience on intergroup bias. *Journal of Personality and Social Psychology, 83*(2), 313–329. https://doi.org/10.1037/0022-3514.83.2.313.

Feng, J., Lazar, J., & Preece, J. (2004). Empathy and online interpersonal trust: A fragile relationship. *Behavior & Information Technology, 23*(2), 97–106. https://doi.org/10.1080/01449290310001659240.

Fogg, B. J., Soohoo, C., Danielson, D., Marable, L., Stanford, T., & Tauber, E. R. (2002). How do users evaluate the credibility of Web sites? In *Proceedings of the 2003 Conference on Designing for User Experiences* (pp. 1–15). https://doi.org/10.1145/997078.997097.

Fung, R., & Lee, M. (1999). EC-trust (trust in e-commerce): Exploring the antecedent factors. In W. D. Haseman & D. L. Nazareth (Eds.), *Proceedings of the Fifth Americas Conference on Information Systems* (pp. 517–519). Milwaukee: Omnipress.

Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation*. Oxford, UK: Academic Press.

Guéguen, N. (2009). Mimicry and seduction: An evaluation in a courtship context. *Social Influence, 4*(4), 249–255. https://doi.org/10.1080/15534510802628173.

Guinote, A. (2007). Power affects basic cognition: Increased attentional inhibition and flexibility. *Journal of Experimental Social Psychology, 43*(5), 685–697. https://doi.org/10.1016/j.jesp.2006.06.008.

Hoeken, H., & Sinkeldam, J. (2014). The role of identification and perception of just outcome in evoking emotions in narrative persuasion. *Journal of Communiation, 64,* 935–955. https://doi.org/10.1111/jcom.12114.

Jacob, C., Gueguen, N., Martin, A., & Boulbry, G. (2011). Retail salespeople's mimicry of customers: Effects on consumer behavior. *Journal of Retailing and Consumer Services, 18,* 381–388.

Jourard, S. M. (1971). *Self-disclosure: An experimental analysis of the transparent self.* Oxford, UK: Wiley.

Jourard, S. M., & Lasakow, P. (1958). Some factors in self-disclosure. *The Journal of Abnormal and Social Psychology, 56*(1), 91–98. https://doi.org/10.1037/h0043357.

Kahneman, D. (2011). *Thinking, fast and slow*. London, UK: Penguin.

Kerr, M., Stattin, H., & Trost, K. (1999). To know you is to trust you: Parents trust is rooted in child disclosure of information. *Journal of Adolescence, 22,* 737–752.

Lee, J., & Soberon-Ferrer, H. (1997). Consumer vulnerability to fraud: Influencing factors. *Journal of Consumer Affairs, 31*(1), 70–89. https://doi.org/10.1111/j.1745-6606.1997.tb00827.x.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology, 46*(6), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025.

Levine, T. R. (2014). Truth-default theory: A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392. https://doi.org/10.1177/0261927X14535916.

Luhmann, N. (1979). *Trust and Power*. New York: Wiley.

Maddux, W. W., Mullen, E., & Galinsky, A. D. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology, 44*(2), 461–468. https://doi.org/10.1016/j.jesp.2007.02.003.

Mason, M. F., Tatkow, E. P., & Macrae, C. N. (2005). The look of love: Gaze shifts and person perception. *Psychological Science, 16*(3), 236–239. https://doi.org/10.1111/j.0956-7976.2005.00809.x.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*(5), 469–480. https://doi.org/10.1023/a:1020278620751.

Miller, N. (2002). Personalization and the promise of contact theory. *Journal of Social Issues, 58,* 387–410. https://doi.org/10.1111/1540-4560.00267.

Morris, M., Nadler, J., Kurtzberg, T., & Thompson, L. (2002). Schmooze or lose: Social friction and lubrication in e-Mail negotiations. *Group Dynamics, Theory, Research and Practice, 6*(1), 89–100. https://doi.org/10.1037//1089-2699.6.1.89.

Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us.* New York, NY: Bloomsbury Press/Bloomsbury Publishing.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205.

Petty, R. E., & Mirels, H. L. (1981). Intimacy and scarcity of self-disclosure: Effects on interpersonal attraction for males and females. *Personality and Social Psychology Bulletin, 7,* 493–503. https://doi.org/10.1177/014616728173020.

Pitesa, M., & Thau, S. (2013). Compliant sinners, obstinate saints: How power and self-focus determine the effectiveness of social influences in ethical decision making. *Academy of Management Journal, 56*(3), 635–658. https://doi.org/10.5465/amj.2011.0891.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49*(1), 95–112. https://doi.org/10.1037/0022-3514.49.1.95.

Ring, P. S., & Van de Ven, A. H. (1992). Structuring cooperative relationships between organisations. *Strategic Management Journal, 13*(7), 483–498. https://doi.org/10.1002/smj.4250130702.

Robbins, J. M., & Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review, 9,* 32–47. https://doi.org/10.1207/s15327957pspr0901_3.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35*(4), 651–665. https://doi.org/10.1111/j.1467-6494.1967.tb01454.x.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23,* 393–404. https://doi.org/10.5465/AMR.1998.926617.

Scissors, L. E., Gill, A. J., & Gergle, D. (2008). Linguistic mimicry and trust in text-based CMC. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (pp. 277–280), San Diego, CA. https://doi.org/10.1145/1460563.1460608.

Scissors, L. E., Gill, A., J, Geraghty, K., & Gergle, D. (2009). In CMC we trust: The role of similarity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 527–536). Boston, MA. https://doi.org/10.1145/1518701.1518783.

Sillence, E., Briggs, P., Harris, P., & Fishwick, L. (2006). A framework for understanding trust factors in web based health advice. *International Journal of Human Computer Studies, 64,* 697–713. https://doi.org/10.1016/j.ijhcs.2006.02.007.

Stajano, F., & Wilson, P. (2011). Understanding scam victims: Seven principles for systems security. *Communications of the ACM, 54*(3), 70–75. https://doi.org/10.1145/1897852.1897872.

Stel, M., & Harinck, F. (2011). Being mimicked makes you a prosocial voter. *Experimental Psychology, 58*(1), 79–84. https://doi.org/10.1027/1618-3169/a000070.

Stel, M., & Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimickees and their interaction. *British Journal of Psychology, 101,* 311–323. https://doi.org/10.1348/000712609X465424.

Street, C. N. H. (2015). ALIED: Humans as adaptive lie detectors. *Journal of Applied Research in Memory and Cognition, 4*(4), 335–343. https://doi.org/10.1016/j.jarmac.2015.06.002.

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology and Behavior, 7*(3), 321–326. https://doi.org/10.1089/1094931041291295.

Swaab, R. I., Maddux, W. W., & Sinaceur, M. (2011). Early words that work: When and how virtual linguistic mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology, 47*(3), 616–621. https://doi.org/10.1016/j.jesp.2011.01.005.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24–54. https://doi.org/10.1177/0261927X09351676.

Thiebach, M., Mayweg-Paus, E., & Jucks, R. (2015). "Probably true" says the expert: How two types of lexical hedges influence students' evaluation of scientificness. *European Journal of Psychology of Education, 30*(3), 369–384. https://doi.org/10.1007/s10212-014-0243-4.

Turner, R. N., Hewstone, M., & Voci, A. (2007). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. *Journal of Personality and Social Psychology, 93*(3), 369–388. https://doi.org/10.1037/0022-3514.93.3.369.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185,* 1124–1134. https://doi.org/10.1126/science.185.4157.1124.

van Baaren, R. B., Holland, R. W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology, 39*(4), 393–398. https://doi.org/10.1016/S0022-1031(03)00014-3.

Vishwanath, A. (2015). Habitual Facebook use and its impact on getting deceived on social media. *Journal of Computer-Mediated Communication, 20,* 83–98. https://doi.org/10.1111/jcc4.12100.

Vishwanath, A., Harrison, B., & Ng, Y. J. (2016). Suspicion, cognition, and automaticity model of phishing susceptibility. *Communication Research, 45*(8), 1146–1166. https://doi.org/10.1177/0093650215627483.

Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems, 51,* 576–586. https://doi.org/10.1016/j.dss.2011.03.002.

Webb, J. T. (1969). Subject speech rates as a function of interviewer behavior. *Language and Speech, 12*(1), 54–67. https://doi.org/10.1177/002383096901200105.

Whitty, M. T. (2015). Anatomy of the online dating romance scam. *Security Journal, 28*(4), 443–455. https://doi.org/10.1057/sj.2012.57.

Williams, E. J., Beardmore, A., & Joinson, A. N. (2017). Individual differences in susceptibility to online influence: A theoretical review. *Computers in Human Behavior, 72,* 412–421. https://doi.org/10.1016/j.chb.2017.03.002.

Williams, E. J., Morgan, P., & Joinson, A. N. (2017). Press accept to update now: Individual differences in susceptibility to malevolent interruptions. *Decision Support Systems, 96,* 119–129. https://doi.org/10.1016/j.dss.2017.02.014.

Workman, M. (2008). Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology, 59*(4), 662–674. https://doi.org/10.1002/asi.20779.

Worthy, C. B., Gary, A. L., & Kahn, G. M. (1969). Self-disclosure as an exchange process. *Journal of Personality and Social Psychology, 13,* 59–63. https://doi.org/10.1037/h0027990.

Why Methods Matter:
Approaches to the Study of Deception
and Considerations for the Future

Zachary M. Carr, Anne Solbu and Mark G. Frank

Humans have intentionally deceived each other for thousands of years (Baron-Cohen, 1999; Bond & Robinson, 1988), yet our understanding of it still remains nebulous. As a whole, researchers have uncovered no *Pinocchio response* (Ekman, 1985; Frank, 2005; Zuckerman, DePaulo, & Rosenthal, 1981), that is, not one deceptive cue[1] that is indicative of deception. However, scientists continue to examine whether any clues, or combination of clues, may in fact be associated with deception (e.g., Frank & Svetieva, 2013). The consideration of so many potential variables presents two concerns for deception researchers: (a) the careful documentation of all steps involved in the research process and (b) building a line of productive inquiry and collaboration with fellow scientists. Only with such transparency and humility will the field of deception advance.

This chapter offers a few considerations that will provide deception researchers with a guided checklist to reference in their conceptualization, operationalization, and design of future studies. To provide context, we outline previous methodological approaches, including cognitive, physiological,

Z. M. Carr (✉) · A. Solbu · M. G. Frank
University at Buffalo, Buffalo, NY, USA
e-mail: zmcarr@buffalo.edu

A. Solbu
e-mail: as255@buffalo.edu

M. G. Frank
e-mail: mfrank83@buffalo.edu

and emotional components of deception that researchers have previously examined. Further, we underscore the point that understanding deception and its associated behavioral manifestations is a multifaceted, diverse, and complex process. Consequently, this was why deception researchers have employed drastically different techniques in their attempts to identify reliable and valid behavioral cues to deceit. Although such variety is to be celebrated for capturing ecological validity, it can also lead to erroneous judgments concerning which cues to deception are potentially more reliable, valid, or generalizable. This applies equally to trying to understand deception from the perspective of both the sender (i.e., encoder) and receiver (i.e., decoder); for the former, to identify which cues may be the most reliable and valid indicators of deceit and for the latter, which cues may be associated with deception detection accuracy or inaccuracy. Throughout the process of laying out these considerations, we hope we can provide scholars the opportunity to connect and bridge empirical gaps, as well as identify new ones.

Initial Approaches to Understanding and Dealing with Deception

Although the formalization of deception as a field of empirical inquiry is relatively young, the observation of its behavioral manifestations is not. The allure of studying deception is, in part, directly driven by society's desire to punish those caught deceiving. As such, past lie catching strategies, crude by our current standard, were employed with draconian implications for the accused (Trovillo, 1939). In the Hindu Vedas written in approximately 900 BC, there are surprisingly detailed instructions of how to detect societal deviants or criminals (e.g., someone who poisoned another) based on their behavior alone. The following illustrates one of the earliest examples of behavioral observations in a deceptive context: "A person who gives poison may be recognized. He does not answer questions, or they are evasive answers; he speaks nonsense, rubs the great toe along the ground, and shivers; his face is discolored; he rubs the roots of the hair with his fingers; and he tries by every means to leave the house…" (Wise, 1845, p. 394; cited by Frank & Svetieva, 2013, p. 472). Similarly, by the time of the ancient Greeks, various physiological signals such as pulse rate, and sweating, were purported to be indicators of deception (Frank & Svetieva, 2013).

Trovillo (1939) suggests that since the time of the ancient Hindus and Greeks, many of the deception detection paradigms were informal or anecdotal. Formalized deception research did not emerge until the nineteenth century with attempts to examine the physiological components of deception, which is unsurprising considering the history described above. As expected, this led to the first efforts to measure deception clues through devices that measured physiological responses (e.g., blood pressure monitors). Trovillo (1939) credits Italian physiologist Mosso (1875) with compiling one of the

first lists of physical behaviors that would subsequently become the standard for deception detection techniques of that time.

Trovillo (1939) further noted that Mosso's most significant observations came from his work on the emotion of fear and its role in deception. Mosso also explored the roles that liars and lie catchers play and how emotions influence those roles. Other researchers expanded upon Mosso's list of physiological indicators to include both verbal and nonverbal cues, for the purposes of observing witnesses giving potentially deceptive testimony. By the end of the nineteenth century, scholars now had something resembling a systematic scheme to classify liars' behavior based on behavioral cues, both verbal and nonverbal; this included descriptions of types of liars, such as those who were flippant, dogged, nervous, humorous, cunning, canting hypocrites, and positive.

Early Twentieth-Century Approaches

Early empirical studies on the psychology of deception employed a diverse set of scenarios to examine the phenomenon (Trovillo, 1939). These scenarios tended to utilize recognition tasks, or *orienting response* tasks. In such tasks, participants were presented with a group of words, either orally or visually. Participants would then be instructed to verbalize the first concept or idea they associated with the stimulus word. Some of these words were associated with the details of a specific crime, others unrelated to the crime. The logic was that those words associated with the crime would be recognized by the perpetrator of the crime, but not by the innocent, triggering the orienting response by the guilty (Waid & Orne, 1982). This is not dissimilar to the reaction when a person hears their name—the body and mind orient to it.

The logic of this paradigm was that the orienting response would trigger physically observable signs of recognition and discomfort (e.g., uncoordinated physical movements, delayed reaction time, repetitions of stimulus words). Further efforts were made in Europe and the US to improve upon the word association test as a primary means of assessing guilt. The use of the word association task continued to be a favored scenario among early deception scholars of the twentieth century and was expanded to include reaction times and more specific criteria for understanding the role of guilt in deceptive contexts (Langfeld, 1920).

An alternative methodological approach focused on the measurement of physiological responses in deceptive contexts (Trovillo, 1939). Some researchers focused on blood vessel dilation and recorded volume changes in water or air devices that were connected to an actuating or recording device. Others examined the constriction of blood vessels and utilized devices akin to our modern-day blood pressure monitors, later including heat readings of blood circulation (Whitehorn, Kaufman, & Thomas, 1935). This then dovetailed with the study of the emotion of fear as envisioned by Mosso (1875)

and its physiological consequences, resulting in early studies examining the simultaneous effects of sudden fear on heart, respiratory, and electrodermal reactions (e.g., galvanic skin response). These different physiological channels each produced their own chart, or graph, hence the many graphs became known as the "poly" graphs. The development of a device to measure them all at once led to the creation of the first polygraph, or "lie detector" machine, thereby hoisting physiological cues atop the behavioral hierarchy as the prominent clues purported to be associated with deception (Larson, 1921). It was this device that was then developed further in both laboratory and naturalistic observations. The polygraph is still utilized today, although both the technological capabilities and practical applications have dramatically increased (Blatz, 1925; Frank & Svetieva, 2013). In sum, the early twentieth-century methodological practices approached deception from a physiological vantage point. However, there were limits to this approach, as emotions can be triggered for reasons other than deception (Ekman, 1985/2001), thus deception researchers sought out other ways to detect behavioral cues to deception.

Later Twentieth-Century Approaches

Categorizing Deceptive Cues in the Body and Face

In 1969, Ekman and Friesen published one of the first influential studies on the observation of potential cues to deception that centered on nonverbal manifestations (e.g., body movements). In essence, they extended previous work on emotion and applied it to deception. Specifically, Ekman and Friesen offered two new categories of cues to be considered—leakage cues and deception cues (Ekman & Friesen, 1969). *Leakage cues*, according to Ekman and Friesen (1969), are clues to the true emotional state of the person that he or she is attempting to conceal (e.g., an angry person *leaking* signs of the emotion of anger through narrowed lips while claiming they are in fact feeling happy). *Deception cues* are those cues that indicate a deception is occurring, although they do not reveal the nature of the information withheld. For example, a liar may commit speech errors as he or she attempts to articulate the lie; these errors are cues that the liar is engaged in cognitive effort to manufacture the falsehood.

The Complexity of Deception

By the early–mid-1980s, researchers postulated that no single behavior or set of behaviors would be uniquely present in deception (Ekman, 1985/2001; Zuckerman et al., 1981). Moreover, agreement emerged that those behaviors that did show a relationship to deception could be caused by reasons other than deception; for example, the nervousness caused by a liar being afraid of getting caught may look identical to the nervousness of a truth teller

who is afraid of being disbelieved (Ekman, 1985/2001). Pop culture, however, is not constrained by the tenets of good science thus there is no shortage of articles claiming that there are easily identifiable cues to spotting deception. For example, Dr. Phil, the famous TV psychologist, claims that he can identify liars based on a single, behavioral cue.[2] He states that if someone touches or rubs their nose during their response to a question, "[t]hat's a dead giveaway" to lying. If this happened to be a ubiquitously reliable sign of lying, this would settle a lot of controversies in the field of deception, making this a much shorter chapter.

The behavioral complexity of deception is acknowledged in an influential review by Zuckerman et al. (1981). They urged deception scholars to consider all the possible psychological or physiological processes that may occur more frequently in liars compared to truth tellers. In fact, they argued that any deception cue we can detect is a result of one (or more) of four different processes—those related to arousal, emotion (i.e., affect), cognition, and behavioral control. Specifically, they argued that (1) *arousal* is generated through activation of the autonomic nervous system, due to emotional or orienting responses, resulting in liars producing increased blinking, pupil dilation, or speech errors; (2) *affect* is generated through the activation of emotions and their resulting expressions of emotions such as guilt or fear in the face, body, or voice; (3) *cognition* is generated through the extra mental effort required to concoct a lie, which creates a higher cognitive load, which is reflected, for example, by an increase in speech errors and increased response latency; and (4) *behavioral control* is generated when a liar *attempts* to inhibit certain behaviors (e.g., reducing hand movements).

The aforementioned processes were thought to be the primary drivers behind behaviors associated with deception, although all of them could be generated for reasons other than deception. However, they did provide the initial scaffolding upon which ensuing scholars attempted to place their models.

An Emotion-Centric Example

Ekman's model had a stronger emotion focus than other models, likely due to his lie scenario which asked participants to lie about what they were feeling as they watched films designed to elicit negative emotions such as distress and disgust (Ekman & Friesen, 1974). He then used the principles of emotion expression as originally articulated by Darwin to explain how facial expressions of emotion are involuntary, thus can "leak" despite efforts to control them (Darwin & Prodger, 1998). The signs of these genuinely felt emotions that were trying to be concealed, like the raised upper lip found in the expression of disgust, Ekman called "leakage cues" (Ekman, 1985/2001). Other cues, like those expressed through the style of speech, such as longer latencies, or increased speech errors, he called "deception cues." He then limited the family of cues associated with deception from the four categories used by

Zuckerman et al. ([1981]) to two categories: *feeling clues* (i.e., emotional) and *thinking clues* (i.e., cognitive).

Ekman's focus on emotion and deception elicited a renewed interest in examining emotion, its subsequent physiological manifestations, and their interactions in deceptive contexts. Ekman hypothesized that the extent to which emotions are elicited when lying is the degree to which the signs associated with emotion in the face (e.g., raised eyebrows pulled together associated with the facial expression of fear) or voice (e.g., raised voice pitch) can accurately predict that a person is lying. What this means is that if emotions are not elicited, then emotional signs will likely not be present. Thus, studies featuring high stakes for the liars—being threatened with 110-decibel blasts of white noise while being detained for an hour—are more likely to elicit fear, distress, or other negative emotions than studies that do not feature and rewards or punishments (Frank & Ekman, [1997]). In this situation, the evaluation of a liar's truthfulness is contingent upon the lie catcher detecting these cues, and then interpreting them properly to render an accurate judgment of veracity (Ekman, [1985/2001]).

Similarly, the liar who experiences pleasure or joy at fooling someone may feel the emotion of enjoyment (Ekman, [1985/2001]). Ekman cites the example of Adolf Hitler's lies about Germany's intentions to British Prime Minister Neville Chamberlain in the beginning of the Second World War. The lie gave Hitler such delight that his secretary reported that, due to his exuberance at successfully perpetrating the lie, Hitler needed to excuse himself from the meeting so he could revel in his accomplishment, even dancing a jig, before composing himself and re-entering the room with Chamberlain. In this particular context, we could predict that Hitler's elation would most likely manifest as a brief smile, appearing to not fit the context or subject matter being discussed in the meeting. It would have been a leakage cue that he was really feeling pleasure as he claimed Germany had no further territorial ambitions. This quick smile of course would not guarantee that a lie was told, but instead its anomalous occurrence could better be understood as a *hot spot* (Ekman, [1985/2001]; Frank, Yarbrough, & Ekman, [2006]), which would alert the astute observer that something was amiss. In fact, identifying these behavioral clues is most often the trigger that initiates a search for *hard evidence* to prove someone has lied (Novotny et al., [2018]).

A Cognition-Centric Example

Cognitions in deception are not just about the mental effort a liar engages to produce a coherent story, but can also be about natural human memory processes (Yuille, [1989]). Credibility analysis involves examining the words spoken by individuals to assess whether they had truly experienced an event, or are in fact fabricating that event. It is based upon what's been called *the Undeutsch hypothesis*, named after German Psychologist Udo Undeutsch,

who proposed that the principles of narrative memory predict that accounts provided by people who have actually experienced an event are qualitatively and quantitatively different from accounts provided by people who have not actually experienced that event (Khoenken, 1989). Thus, someone who truly experienced an event will produce an account that is physically possible, contains the context surrounding the event, will reference interactions with others, will have levels of detail appropriate to how significant the event is to the person (more significant equals more detail, and vice versa), and where the person can more readily recount the event starting at different points in the narrative (Yuille, 1989). Likewise, Johnson and Raye (1981) proposed a *reality monitoring model* based upon that assumption, where experienced events are more likely to feature sensory information, or awareness of one's own thoughts and thinking processes (i.e., cognitive operations). Other cognitive models suggest that the words chosen by liars will be such as to psychologically separate themselves from the act; for example, dropping the person pronouns of "I" or "me," not referring to people by specific names, using more passive voice, employing more hedges, and so forth (Newman, Pennebaker, Berry, & Richards, 2003). Thus, a truth teller might say "Zach and I drove to the store and I bought a pack of gum," whereas a liar might say "Went to the store, for the most part only bought some gum."

A Communicative-Centric Example

Up until the late 1990s, most of the perspectives and methodologies in deception research originated from the psychological discipline, which focused on understanding internal processes associated with deception and their related behaviors. However, Buller and Burgoon (1996) offered a communicative perspective that highlighted interpersonal communication processes and their role in deception—*interpersonal deception theory*. Their main argument was that when someone is engaged in active deception, they are also simultaneously engaged in several tasks (e.g., conveying their message, monitoring their target's perception of the message). Buller and Burgoon maintained that over time, the deceiver would gain feedback and information that would allow them to perpetrate a more convincing lie. Such feedback could put the liar at ease and would cause the liar to display specific behaviors associated with this increased ease of lying (e.g., increased nonverbal immediacy).

Buller and Burgoon argued that a main moderating variable was the deceiver's motivation. Given the complex taxonomy of motivation in previous research (e.g., Metts, 1989; Turner, Edgley, & Olmstead, 1975), effectively defining and creating motivation is difficult. Buller and Burgoon (1996) proposed three different types of motivations: *instrumental*, which is related to the task; *relational*, which is related to the relationship status of participant and interviewer; and *identity*, which is related to how the task reflects upon how the participant views him or herself. Depending upon which motivation

was driving the deceiver's actions, specific behavioral cues could be predicted. What is new here is that deception is acknowledged as being a complex and intricate process, with the primary focus having shifted to the overall communicative exchanges that take place between both sender and receiver in a deceptive context. However, this work was critiqued by psychologists who stated it did not make testable predictions, and that some of the analyses were confounded along with some other methodological issues (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996).

Another communicative approach aimed at the measurement of the multiple behaviors occurring within an interaction, with a specific focus on the nonverbal *synchrony* of all communicators. Synchrony is defined as the "similarity in rhythmic qualities and enmeshing or coordination of the behavioral patterns of both parties" (Burgoon, Stern, & Dillman, 1995, p. 128), and typically has been studied in body language, linguistics, paralinguistics, gaze, and facial affect (Cappella, 1990, 1991, 1997). For instance, Driskell, Salas, and Driskell (2012) investigated behavioral synchrony in dyads of truthful and lying police officers and firefighters standing next to each other, while being interviewed by an experimenter. They found that those dyads where both were truthful exhibited greater synchrony with each other in terms of mutual gaze and speech than those who were lying.

Taken together, the emotion-, cognition-, and communicative-centric examples reveal that scholars often view phenomena through the lens of their own expertise—the people who study emotion focus more on emotional reactions; those who study cognition and memory focus more on word choice in the account provided, and those who study communication focus more on back and forth interaction of senders and receivers. It would be a mistake to consider just one perspective to be correct, or that one approach fully explains the phenomenon of human deception. Weaving these approaches together is most likely the best approach. Otherwise, as scholars we get into a situation akin to the parable of the seven blind men feeling the elephant, each describing a very different animal because of their attention to just one area of the elephant (e.g., one man describing an elephant as long and muscular versus another man describing an elephant as wide and flat). To fully understand deception, which is at the crossroads of emotion, cognition, behavioral management, and social interaction, we need to understand all its various components, and how they fit together. With that in mind, there are methodological issues that cut across all perspectives that we believe researchers and relevant professionals can benefit from considering.

CONSIDERATIONS FOR STUDYING DECEPTION

The field of deception is characterized by a myriad of variables, foci, and approaches, but appears to congeal around three general processes driving the behavioral clues that may be related to deceit—those produced by *cognition* such as mental effort and narrative memory processes; *emotion*, which

subsumes arousal and its associated physiological responses; or *behavioral control*, such as attempts to manage those cues or signs; or some combination therein (Frank & Svetieva, 2013; Zuckerman et al., 1981). Such diversity of processes in deception is not problematic; rather, its complexity should be relished, as very few behavioral phenomena can be understood and examined from such a wide range of academic foci. The intersection of these three processes is the nexus of what humans wrestle with in every interpersonal situation we encounter because they all require us to think, feel, and manage our impressions. Thus, we suggest the best way to study this inherent complexity is to systematically consider a number of issues when designing any deception study.

Consideration #1: Check the Definitions

The deception literature is replete with examples of empirical papers that introduce a concept but then does not define it. Such a problem may appear to be trivial at first glance, but conceptual definitions are the building blocks of any valid operationalization, and for example, ultimately influence any experimental design. If such conceptualizing remains ambiguous, a researcher may not know exactly what it was that he or she has discovered. This can confuse researchers who may attempt to study similar constructs or variables, but in reality are studying different phenomena. Take for example how various deception researchers conceptualize and define *arousal* (Frank, 2005). Multiple empirical studies have examined arousal, yet they fail to offer an explicit definition of it. In fact, we see arousal represented by such varying constructs as an intense emotional response (e.g., fear) (Frank, 1989), an orienting response (deTurck & Miller, 1985), anger, or sometimes it has been understood more ambiguously, with multiple variables of interest included (Burgoon, Buller, Hale, & de Turck, 1984; see discussion by Waid & Orne, 1982).

A similar problem is found when defining a smile. A review of the research examining the role of smiling in deception showed that rarely did researchers define what they meant by a smile (Frank, 2002). This turns out to be important, as not all smiles are the same, and some types of smiles—masking smiles, used to conceal a different emotion—are more implicated in deception than others (Ekman, O'Sullivan, & Friesen, 1988). Taken together, this means if researchers are going to study a phenomenon, or challenge it, they must explicitly define what it is they are studying, and what is the criterion or criteria that represent it. Only then will we know if the subsequent results clarify or obfuscate our understanding of deception (Gendron, Roberson, van der Vyver, & Barrett, 2014).

The above problem may appear to be innocuous, but to restate the argument proposed by Frank (2005), the lack of explicit definitions is a significant impediment to advancing the field of deception. This is akin to a researcher claiming to study nonverbal communication but examining *only* facial

expressions without defining nonverbal communication; whereas another claims to study nonverbal communication and examines *only* gestures, again without defining nonverbal communication. Both researchers inevitably would reach drastically different conclusions about the utility of nonverbal communication because both were studying completely different constructs while claiming to study the same one. Although this is a hyperbolic example, it vividly illustrates a very real issue for the field of deception.

Therefore, we strongly urge all deception researchers to explicitly define their respective constructs of interest. This applies to broader variables (e.g., arousal), but it also applies to more specific constructs (e.g., the facial muscles that define a smile). By delineating a clear, systematic definition of terms and conceptualizations, deception researchers (and reviewers alike) can appropriately and efficiently assess the validity of a particular study's findings and generalizability concerning the particular set of behavioral cues studied. Further, this enables more effective comparisons between different deception scenarios—like a false opinion compared to a fake theft compared to a cheating task—and how well the features of the laboratory map onto the features of real-world deception situations.

Ironically, the definition of deception itself has undergone a number of varying conceptualizations (Burgoon & Buller, 1994; Miller & Stiff, 1993)—all with unique preferences as to the primary variable of interest that comprises deception. So, in taking our own advice, and for the purposes of this chapter we adhere to the comprehensive definition of a lie as employed by Ekman (1985/2001)—a deliberate attempt to mislead, without prior notification of the target of the lie. There are other definitions, such as that by Masip, Garrido, and Herrero (2004): "Deception is defined as the deliberate attempt, whether unsuccessful or not, to conceal, fabricate, and/or manipulate in any other way, factual and/or emotional information, by verbal and/or nonverbal means, in order to create or maintain in another or others a belief that the communicator himself or herself considers false" (p. 148). It is longer than Ekman's definition and articulates more, but differs in whether prior notification is important. Ekman would not consider a poker bluff a lie, as bluffing is part of the game, thus the opponent understands that this may happen. It would be deception, but not a lie, according to Ekman (1985/2001). Philosophers such as Bok (1980) have articulated deception as being a higher order category of behavior, done deliberate or not, compared to a lie, which is a subcategory of deception that is always done deliberately. To this aspect of deliberateness, the Ekman (1985/2001) and Masip et al. (2004) definitions agree.

Consideration #2: Attend to the Stakes Involved

Our second point builds on the previous consideration by urging deception researchers to carefully consider the stakes associated with the lie. The stakes refer to the pattern of incentives or punishments for successful or unsuccessful

lying or truth telling. A police interrogation is a very high-stakes situation—an unsuccessful lie can result in a person going to jail; a successful lie, in some instances, can result in a person literally getting away with murder. Ethical constraints make it impossible to employ and study these very high-stakes situations in the laboratory. The majority of the published research literature features lies with little to no stakes (Frank & Svetieva, 2013). This is beyond definitional convenience, but will affect the behaviors shown by liars versus truth tellers, because the higher stakes would raise the emotional reactivity of the liar, and in turn elicit emotional clues. For example, meta-analyses showed that across all studies, the effect size of behaviors associated with emotions, such as *nervousness,* was significant but weak (DePaulo et al., 2003). However, when the studies in the meta-analysis were divided into those with high motivation for participants, and those with low motivation, the effect size for the former now jumped to being moderate in strength, whereas the effect size for the latter was functionally zero (Frank & Svetieva, 2013 based upon DePaulo et al., 2003). Stakes would also have huge implications for ecological validity and hence may limit our abilities to generalize results from low-stake experiments to real-world law enforcement situations. Stakes can affect not just emotions, but also cognitive activity, and our abilities to control our expressive behavior in general. As such, researchers should consider stakes when applying research results to real-world situations.

One could not be faulted then to consider using real-world material to study high-stakes lies (e.g., Mann, Vrij, & Bull, 2002; ten Brinke, Stimson, Carney, 2014). However, the downside of this is the loss of control over knowing the ground truth (Frank, 2005). What is *true* and *false* may be easier to determine than what a subject believes to be true, or believes to be false. Witnesses can be inaccurate without lying. Subjects misremember events. How often have people discussed a specific event with friends, just to find out not all agree on exactly what happened? What this all suggests is that the criterion for ground truth is not as clear as it seems. In the laboratory, we can control the situation we are testing, and check to see if the participant in our study stole the ring or changed their opinions, and thus we can know for sure that they are lying. In the real world, we have to choose a defensible criterion or criteria for what is the ground truth (reality) because the obvious is not so obvious. For example, if examining a suspect interview by the police, we could argue that the criterion for ground truth would be conviction by a jury. Or better yet, we may argue the criterion for ground truth is the suspect's confession. However, this is more complicated than it seems, as newspapers routinely run articles showing that people are falsely convicted, and people confess falsely. Moreover, a researcher may want to have a truth from a suspect in a real-world interrogation so as to compare behaviors with what is believed to be the suspect's lie. Researchers again must clarify the criteria for truth. Is it the suspect's admission of guilt for a prior conviction? Is it information that can be clearly verified, such as where they receive their mail? Regardless, we suggest researchers articulate their criteria for

everything—stakes, ground truth, and so forth, very carefully and be aware of the limits of those decisions in their studies.

Not all the lies we study are high stakes; many are considered insignificant or what is colloquially known as the *white lie* (DePaulo et al., 1996). Such lies are often told in everyday occurrences as they act as *social lubrication* and serve a valued politeness function. One communicator may ask another, "How you are doing?", and the acceptable response is to say something akin to, "I am fine, how are you?" However, if you are not doing fine, you are *still* expected to reply "I am doing fine." If you divulge all of the things that may be going wrong that particular day, the social consequences may be severe, as people may want to avoid you and your monologues. So, given how frequently these white lies occur (e.g., 1–2 times every day; DePaulo et al., 1996), and their value to our overall social functionality, there is little doubt that such lies are worthy of study and observation.

However, the lies that may be important to detect are the high-stakes lies, like those found in law enforcement, intelligence, or counter-terrorism. Scholars can identify the exact stakes and pressures found in real-life counter-terrorism situations and then attempt to incorporate some representation of them in the laboratory setting, even though it would not be ethical to fully recreate those stakes (Frank, 2005). For example, previous deception experiments have utilized incentives for participants who are successful liars (e.g., monetary incentives, favorable perception of intelligence), or by crowning a "best liar," with the intent to increase the motivation for participants to deceive (DePaulo, Lanier, & Davis, 1983; Kraut & Poe, 1980). Fewer studies, however, provide punishments for those who attempted to lie but were unsuccessful (Frank, 2005). Some studies have motivated research participants by telling them if they get caught lying, they will not only lose their payment for being in the study, but they will be detained and subjected to loud blasts of white noise (Frank & Ekman, 1997). All of the above points are contingent upon which type of scenario the deception researcher is interested in generalizing. Extreme caution should be taken to avoid conflating results from high- and low-stakes deception studies. Surprisingly, this is common in the deception field, with very few researchers attempting to differentiate this dichotomy appropriately (e.g., Hartwig & Bond, 2011). Consequently, scholars may be promoting particular behavioral clues for training officers that do not appear in high stakes, real-world settings, or may be failing to train officers on behavioral clues that may be valid in these situations. Not only would officers suffer, but so would public safety in general.

Consideration #3: Decide on Choice or Random Assignment

A key determinant of the ecological validity of a deception experiment is whether or not participants can choose to lie or if they are randomly assigned to tell the truth or lie. In everyday life, we can reasonably assume that people

are not assigned to lie. Rather, we choose whether to lie. When we choose to lie, we *own* our lie. This may give us added motivation to successfully perpetrate the lie, but also may have resultant effects on making us feel more guilt or shame, or lower our self-esteem or self-image (DePaulo et al., 2003; Frank, 2005).

The notion of choosing to lie inherently clashes with standard scientific experimental practices, where researchers must randomly assign participants to conditions in order to infer causality. Random assignment does raise the internal validity of a study, but may decrease its external validity because a lie is rarely assigned in real life. Therefore, the deception researcher must wrestle with choosing to represent a more ecologically valid paradigm (e.g., allowing participants to choose whether or not to lie) or one that is more internally valid (e.g., randomly assigning participants to lie). However, the lack of random assignment is not a death knell for the experiment. It merely changes what may have been a true experimental design into a quasi-experimental design. Without quasi-experimental designs, scholars could not study smokers versus nonsmokers or men versus women. But what the quasi-experimental design does is hinder one's ability to say that a participant's lying *caused* a certain behavior in the participant.

An additional problem created by allowing choice is that the pattern of incentives in the study of lying and truth telling may be so that only 5% of the participants choose to lie. In this case, a researcher can adjust the incentives for lying or truth telling in order to obtain an approximate 50/50 split for participants choosing lying or truth telling to facilitate analysis. Sometimes, in order to obtain an appropriate number of liars or truth tellers, over-recruitment is necessary, because it is an assumption for many inferential statistics that comparisons between groups stay within a 3–1 ratio as major imbalances in comparison sizes affect assumptions about homogeneity of variances (Keppel, 1991). Regardless, one must still exercise caution when interpreting results concerning what percentage choose to tell the truth or to lie, as it was the researcher who created the pattern of incentives to obtain the approximate 50/50 split.

Consideration #4: Decide Whether to Use Sanctioned vs. Unsanctioned Lies

An *unsanctioned lie* is one in which a participant is *not* told that lying is part of the experiment (Feeley & deTurck, 1998). In the laboratory setting, the majority of lies observed are *sanctioned*, meaning that participants are expected and authorized to lie, or are explicitly told to lie by the experimenter (usually after being randomly assigned to do so). Although sanctioned lies are widely used, they do have drawbacks. Sanctioned lies may produce an unnatural response by a participant. For instance, a participant may not want to lie, be uncomfortable lying, or be a really horrible liar, but because of random assignment, they are expected to lie. If the lie is sanctioned, they may not display their usual behaviors when lying, because the authorization

to lie removes some of the participant's responsibility for their actions, thus reducing feelings of guilt. Or a participant may feel less motivation to lie successfully because they know that there will be no consequences for lying. The result may be that the participant will exhibit behavioral cues that may be more or less explicit than what would be observed from them in an unsanctioned lie told outside of the laboratory.

One may argue that allowing subjects to choose whether to lie makes it unsanctioned, even though that choice to lie or not is explicitly sanctioned by the experimenter. Unfortunately, for the experimenter, true unsanctioned lies typically are much harder to implement ethically in the laboratory. The *cheating paradigm* is an example of an unsanctioned lie used in deception research. This situation involves the participant and his or her teammate (usually a confederate) in a contest of some sort (e.g., estimating the number of dots on a sheet of paper). The teammate discovers or steals the correct answer to the contest. Later, the participant is asked how his or her team came up with the correct answer to win the contest—and usually a minority of those participants will lie and not reveal that they won because their teammate stole the answer (e.g., Exline, Thibaut, Hickey, & Gumpert, 1970). Different variations include allowing students to grade their own quizzes and then lie about cheating (Stiff, Corman, Krizek, & Snider, 1994), and using confederates to encourage cheating (e.g., Exline et al., 1970). For instance, Feeley and DeTurck (1998) used a research confederate to induce students to cheat by having the confederate offer participants the answers. The experimenter revealed to half the students that they were aware of the cheating but asked them to conceal it (i.e., sanctioned deception), but did not give any instructions to the other half (i.e., unsanctioned deception). Regardless, one must be careful when studying unsanctioned lies—they may be more ecologically valid, but they often pose issues with internal validity (see Frank, 2005, for a more thorough discussion).

Consideration #5: Consider the Interview Structure

Our final consideration addresses perhaps the most ignored aspect of deception studies—the interview in which the participant tells his or her lies or truths (Frank, 2005; Vrij, Mann, & Fisher, 2006). In the real word, law enforcement officials will engage in interviews with both suspects and witnesses while noting both verbal and nonverbal cues to help further their line of questioning (Blair, Levine, & Shaw, 2010; Frank et al., 2006). In the laboratory, much of the literature has not considered the type of interview (e.g., accusatory or providing a recollection only), the forms of the questions (e.g., open-ended or close-ended), or the relationship between the interviewer and interviewee (e.g., adversarial or friendly), yet all these features are important (Frank, 2005). We now know *how* interviewers ask questions (i.e., order of accusation; e.g., Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2007; Vrij et al., 2006),

or how the interviewer deploys the evidence in his or her possession (Hartwig, Bond, Stromwell, & Kronkvist, 2006), is vitally important as it affects the type and amount of behavioral clues expressed by the potentially deceptive participant. In a real-world investigation, an interviewer can deploy previously unknown information to the interviewee and follow leads without any set form or question protocol, often resulting in very high deception detection accuracy (Levine, 2015). Given that most real-life investigative interviews are not tightly structured, like they are in the laboratory, their replication in the laboratory is difficult (but see Frank & Ekman, 1997; Levine, 2015). Artificially constrained interviews from a laboratory are difficult to extrapolate to the give and take of a real-world investigation, where the investigator can follow up on unclear statements or unusual reactions by the lying individual.

Summary of Considerations

Therefore, the most critical question for any deception researcher interested in designing an ecologically valid deception scenario is "Does my *artificial* interview accurately reflect an interview in the real-world?" The answer to this question is usually a resounding *no*. However, researchers who want to generalize their results should ask, "How *close* can I get my *artificial* interview to reflect a real-world interview?" And then report results and discussions with humility when generalizing those results, as they too will not perfectly match a real-world high-stakes interview. The problem is that there are numerous possible variables to consider—from the details of the account (e.g., liars vs. truth tellers; see Sporer & Schwandt, 2006), to the cognitive demands associated with fabricating a believable account, to the behavioral markers believed to manifest as a result of creating and maintaining a believable account (DePaulo et al., 2003). Given the vast number of interview variables to consider, we do not hesitate to state that attempting to control all of them is nearly impossible. Other considerations include the various structural features of a deception experiment that are not varied within an experiment, but may still affect the results. For example, the novelty of the situation—the location, arrangement of chairs, salience of the cameras—all need to be taken into account. New situations for participants draw more mental effort; familiar ones less. Distractions can affect cognitive processing too; for example, we don't know how the blinking red light on a camera that is recording the interaction may affect participants' behaviors.

The interview is now known to be a strong influence on the results obtained in both the laboratory and the real world, thus must be considered with great care. Researchers should thoughtfully think through how they script their interviews, and carefully consider how much freedom interviewers have to interact with participants (e.g., rapport with interviewee, deviating from script). All of these decisions will affect both the internal and external validity of future studies of deception. Thus, some of the best tools

and processes for scholars are to make sure to note their procedures in detail, define key terms, and explain their decisions, so that other researchers and practitioners can properly interpret results in context.

CONCLUSION

The above considerations offer both general and specific guidance for deception researchers as they conceptualize future studies. Additionally, these considerations provide an evaluative framework for reviewers and professionals alike. Future opportunities and challenges for the deception researcher involve taking advantage of increasingly sophisticated technology that can better measure one's autonomic nervous system, the endocrine system, demeanor, brain activity, and additional behaviors automatically and efficiently. Some results using these technologies have already been reported, such as studies that use remote detection of psychophysiological measures including blood flow and facial temperature (Tsiamyrtzis et al., 2007), or automatic facial expression recognition and analysis (e.g., Bartlett, Littlewort, Frank, & Lee, 2014), body and eye tracking (e.g., Bhaskaran, Nwogu, Frank, & Govnidaraju, 2011), vocal analysis (Levitan et al., 2016), and neural activity measures (Langleben & Moriarty, 2013). Taken together, researchers are presented with opportunities to more easily capture multiple behavioral responses simultaneously. Examples include controllable and conscious behaviors, such as posture and arm movements, as well as less controllable and more unconscious behaviors, such as hormone level and skin conductance.

Furthermore, sophisticated software now allows for the examination of synchrony and interplay in dyadic behavior (see Buller & Burgoon, 1996; Dunbar, Jensen, Tower, & Burgoon, 2014). Coupling this with advanced computer interfaces allows for labeling and visualizations of events in all channels, enabling context-dependent analysis (see Blair et al., 2010; Bartlett et al., 2014). Advanced computer algorithms can then be implemented to find patterns across and within participants, who tell both truths and lies. It is our belief that deception research projects will become more collaborative in the future between behavioral scientists and neuro, cognitive, and computer scientists, and with their resultant technology and its inherent limitations. It is clear from the history of deception research that anytime someone has discovered some new technology, it will be eventually applied to seeing how it helps us detect lies. For example, functional magnetic resonance imaging was developed to observe brain activity and locate lesions; yet this technology was eventually used to study what sections of the brain activate when someone tells a lie (Langleben & Moriarty, 2013). However, what can never be replaced by technology is the use of sound approaches to the study of deception, as discussed in this chapter. The best technology used in a poorly designed study, coupled with unrealistic or unnatural scenarios, will produce misleading results in the tradition of the proverbial *garbage-in, garbage-out*

phenomenon. Instead, as the field of deception research continues to evolve, it is imperative that researchers approach the study of deception in a more united, clearly defined fashion for the betterment of our collective scholarly knowledge and for those professionals who rely on it.

Notes

1. For the purposes of this chapter, we use *cue* and *clue* interchangeably.
2. See Dr. Phil's website and article titled, "How to Spot a Liar" at https://www.drphil.com/advice/how-to-spot-a-liar/.

References

Baron-Cohen, S. (1999). The evolution of a theory of mind. In M. C. Corballis & S. E. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution.* Oxford, UK: Oxford University Press.

Bartlett, M. S., Littlewort, G. C., Frank, M. G., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. *Current Biology, 24*(7), 738–743.

Bhaskaran, N., Nwogu, I., Frank, M. G., & Govindaraju, V. (2011). Deceit detection via online behavioral learning. In *Proceedings of the 2011 ACM Symposium on Applied Computing* (pp. 29–30). ACM.

Blair, J. P., Levine, T. R., & Shaw, A. S. (2010). Content in context improves deception detection accuracy. *Human Communication Research, 36,* 420–439.

Blatz, W. E. (1925). The cardiac, respiratory, and electrical phenomena involved in the emotion of fear. *Journal of Experimental Psychology, 8*(2), 109–132.

Bok, S. (1980). The self deceived. *Social Science Information, 19*(6), 923–935.

Bond, C. F., & Robinson, M. (1988). The evolution of deception. *Journal of Nonverbal Behavior, 12,* 295–307.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18*(2), 155–184.

Burgoon, J. K., Buller, D. B., Hale, J. L., & de Turck, M. A. (1984). Relational messages associated with nonverbal behaviors. *Human Communication Research, 10*(3), 351–378.

Burgoon, J. K., Stern, L. A., & Dillman, L. (1995). *Interpersonal adaptation: Dyadic interaction patterns.* New York, NY: Cambridge University Press.

Cappella, J. N. (1990). On defining conversation coordination and rapport. *Psychological Inquiry, 1,* 303–305.

Cappella, J. N. (1991). Mutual adaptation and relativity of measurement. In B. Montgomery & S. Duck (Eds.), *Studying interpersonal interaction* (pp. 103–117). New York, NY: Guilford.

Cappella, J. N. (1997). Behavioral and judged coordination in adult informal social interactions: Vocal and kinesic indicators. *Journal of Personality and Social Psychology, 72,* 119–131.

Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007). Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling, 4*(3), 167–180.

Darwin, C., & Prodger, P. (1998). *The expression of the emotions in man and animals*. Oxford, UK: Oxford University Press.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979–995.

DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45*(5), 1096–1103.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118.

de Turck, M. A., & Miller, G. R. (1985). Deception and arousal: Isolating the behavioral correlates of deception. *Human Communication Research, 12,* 181–201.

Driskell, J. E., Salas, E., & Driskell, T. (2012). Social indicators of deception. *Human Factors, 54,* 577–588.

Dunbar, N. E., Jensen, M. J., Tower, D. C., & Burgoon, J. K. (2014). Synchronization of nonverbal behaviors in detecting mediated and non-mediated deception. *Journal of Nonverbal Behavior, 38,* 355–376.

Ekman, P. (1985/2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. New York, NY: Norton.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*(1), 49–98.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology, 29*(3), 288–298.

Ekman, P., Friesen, W. V., & O'sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology, 54*(3), 414–420.

Exline, R. V., Thibaut, J., Hickey, C. P., & Gumpert, P. (1970). Visual interaction in relation to machiavellianism and an unethical act. In R. Christie & F. L. Geis (Eds.), *Studies in Machiavellianism* (pp. 53–75). New York, NY: Academic Press.

Feeley, T. H., & Deturck, M. A. (1998). The behavioral correlates of sanctioned and unsanctioned deceptive communication. *Journal of Nonverbal Behavior, 22*(3), 189–204.

Frank, M. G. (1989). Human lie detection ability as a function of the liar's motivation (Unpublished doctoral dissertation), Cornell University, Ithaca, NY.

Frank, M. G. (2002). Smiles, lies, and emotion. In M. H. Abel (Ed.), *An empirical reflection on the smile* (pp. 15–43). Lewiston, NY: Edwin Mellen Press.

Frank, M. G. (2005). Research methods in detecting deception research. In J. Harrigan, R. Rosenthal, & K. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 341–368). Oxford, UK: Oxford University Press.

Frank, M. G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. *Journal of Personality and Social Psychology, 72*(6), 1429–1439.

Frank, M. G., & Svetieva, E. (2013). The role of nonverbal behavior in detecting and telling lies. In M. L. Knapp & J. A. Hall (Eds.), *Nonverbal Communication* (pp. 471–511). New York, NY: Mouton de Gruyter.

Frank, M. G., Yarbrough, J. D., & Ekman, P. (2006). Investigative interviewing and the detection of deception. In T. Williamson (Ed.), *Investigative Interviewing: Rights, Research, Regulation* (pp. 229–255). Abingdon, UK: Routledge.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion, 14*(2), 251–262.

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*(4), 643–659.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior, 30*(5), 603–619.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67–85.

Keppel, G. (1991). *Design and analysis*. Engelwood Cliffs, NJ: Prentice Hall.

Kohnken, G. (1989). Behavioral correlates of statement credibility: Theories, paradigms, and results. In H. Wegener, F. Losel, & J. Haisch (Eds.), *Criminal behavior and the justice system: Psychological perspectives* (pp. 271–289). Berlin, Germany: Springer.

Kraut, R. E., & Poe, D. B. (1980). Behavioral roots of person perception: The deception judgments of customs inspectors and laymen. *Journal of Personality and Social Psychology, 39*(5), 784–798.

Langfeld, H. S. (1920). Psychophysical symptoms of deception. *The Journal of Abnormal Psychology, 15*(5–6), 319–328.

Langleben, D. D., & Moriarty, J. C. (2013). Using brain imaging for lie detection: Where science, law, and policy collide. *Psychology, Public Policy, and Law, 19*(2), 222–234.

Larson, J. A. (1921). Modification of the Marston deception test. *Journal of the American Institute Criminal Law & Criminology, 12,* 390–399.

Levine, T. R. (2015). New and improved accuracy findings in deception detection research. *Current Opinion in Psychology, 6,* 1–5.

Levitan, S. I., An, G., Ma, M., Levitan, R., Rosenberg, A., & Hirschberg, J. (2016). Combining acoustic prosodic, lexical, and phonotactic features for automatic deception detection. In *INTERSPEECH* (pp. 2006–2010).

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies and videotape: An analysis of authentic high-stakes liars. *Law and Human Behavior, 26,* 365–376.

Masip, J., Garrido, E., & Herrero, C. (2004). Defining deception. *Anales de Psicología/Annals of Psychology, 20*(1), 147–172.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6*(2), 159–179.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*. Thousand Oaks, CA: Sage.

Mosso, A. (1875). Sopra un nuovo metodo per scrivere i movimenti dei vasi sanguigni nell'uomo. *Atti della R. Accademia delle Scienze di Torino, 11,* 21–81.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665–675.

Novotny, E., Carr, Z., Frank, M. G., Dietrich, S. B., Shaddock, T., Cardwell, M., & Decker, A. (2018). How people really suspect and discover lies. *Journal of Nonverbal Behavior, 42*(1), 41–52.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20*(4), 421–446.

Stiff, J., Corman, S., Krizek, B., & Snider, E. (1994). Individual differences and changes in nonverbal behavior: Unmasking the changing faces of deception. *Communication Research, 21*(5), 555–581.

ten Brinke, L., Stimson, D., & Carney, D. R. (2014). Some evidence for unconscious lie detection. *Psychological Science, 25*(5), 1098–1105.

Trovillo, P. V. (1939). A history of lie detection. *Journal of Criminal Law and Criminology, 29*(6), 848–881.

Tsiamyrtzis, P., Dowdall, J., Shastri, D., Pavlidis, I. T., Frank, M. G., & Ekman, P. (2007). Imaging facial physiology for the detection of deceit. *International Journal of Computer Vision, 71*(2), 197–214.

Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology, 11*, 69–89.

Vrij, A., Mann, S., & Fisher, R. P. (2006). An empirical test of the behaviour analysis interview. *Law and Human Behavior, 30*(3), 329–345.

Waid, W. M., & Orne, M. T. (1982). The physiological detection of deception: The accuracy of polygraph testing can be affected by such variables as attentiveness, drugs, personality, and the interaction between examiner and subject. *American Scientist, 70*(4), 402–409.

Whitehorn, J. C., Kaufman, M. R., & Thomas, J. M. (1935). Heart rate in relation to emotional disturbances. *Archives of Neurology & Psychiatry, 33*(4), 712–731.

Wise, T. A. (1845). *Commentary on the Hindu system of medicine.* Calcutta, India: Baptist Mission Press.

Yuille, J. C. (1989). *Credibility assessment.* Berlin, Germany: Springer Science & Business Media.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology, 14*, 1–59.

Detecting Deceptive Communication

An Overview of Detecting Deceptive Communication

Timothy R. Levine

The title of the chapter exemplifies truth in advertising. This chapter provides a state of the art overview of theory and research on detecting deceptive behavior. Doubt it? My guess is no, you did not doubt my claim at all. Even if you are now wondering about the honesty of my initial claims, you only did so after I raised the issue. I doubt suspicion about the honesty and accuracy of the first two sentences would have occurred to you had I not asked if you have doubts.

Now that I have raised the issue of truth in advertising, I challenge you to try to ascertain the honesty of my claims for yourself. How might you do that? The best approach is to read this chapter for yourself. See if it really is about detecting deception. If it is, then my opening claim really was honest and accurate. Alternatively, if the chapter either fails to provide an overview or it is not really about detecting deception, then you have caught me in a lie (or at least in making a false statement). Either way you will not only have engaged in an exercise in deception detection (i.e., correctly distinguishing between honest-truthful communication and deceptive communication) but will have also had a demonstration of a couple of the most important take-home messages obtained from three-quarters of a century of social scientific research. Specifically, suspicion requires prompting, and once suspicious of some communication, if you can, seek to objectively fact check it.

Let me boldly make the further claim that more likely than not much of what you think you know about deception detection is probably false.

T. R. Levine (✉)
University of Alabama at Birmingham, Birmingham, AL, USA
e-mail: levinet@uab.edu

As just one example, let me ask you, what's the best way to tell if someone is lying? Surveys find that when asked the open-ended question, "how can you tell someone is lying?", the answers that people most often give have little correspondence with valid lie detection methods. What people think they know about deception detection is the stuff of folk wisdom, not scientific evidence. So, what is the best way to tell someone is lying? The number one most frequent answer to that question on surveys is that a liar won't look you in the eye (Bond & The Global Deception Research Team, 2006). Interestingly, this research finds cross-cultural consensus in this belief. People everywhere think that liars avert their gaze. Yet research also finds that the correlation between gaze aversion and actual honesty-deception is zero (DePaulo et al., 2003). Gaze has no validity as a deception cue. Reliance on eye gaze to detect deception leads to mere chance-level accuracy because eye behavior is not *diagnostic*.

Terminology

Before getting into the theory and research on deception detection, let's be clear about how a few important words are being used here. I'll presume you already have an understanding of the word deception (see Chapter 2 in this handbook by Pamela Kalbfleisch and Tony Docan-Morgan). In this chapter, we will be talking about senders and judges. Senders are the people whose communication is under scrutiny. Senders may be lying or maybe they are honest. Judges are the people trying to detect deception or ascertain honesty.

Deception detection refers to correctly distinguishing between honest and deceptive messages. Think of it like a true–false test. Some statements are true (honest). Some are false (lies). Deception detection involves sorting out which are which.

Deception detection *accuracy* is a score on this true–false test. We take the number right (truths correctly judged as true and lies correctly identified as lies) and divide by the total number of judgments (correct judgments plus errors) to get a proportion or percentage. Being right every time yields 100%. Mistaking all truths for lies and all lies for truths would yield 0%. Getting half right yields 50%. We can call this correct truth–lie discrimination percentage *total accuracy* (or just *accuracy* for short) because it averages across truths and lies.

We can also score just the truths or only the lies. The percent of honest messages judged correctly is *truth accuracy* while *lie accuracy* refers to the percentage of lies judged correctly. It is a good idea to consider truth and lie accuracy separately (in addition to total accuracy) because truth-bias affects truth and lie accuracy differently. The difference between truth accuracy and lie accuracy is called the *veracity effect* (Levine, Park, & McCornack, 1999). *Truth-bias* is the tendency for people to believe others (Levine et al., 1999). People tend to judge others' messages as honest more often than otherwise, and thus people are typically truth-biased (Levine et al., 1999). When people

are truth-biased, truth accuracy is higher than lie accuracy. This is the *veracity effect*. The veracity of the message (if it is honest or a lie) affects accuracy (if a judge is correct). When people judge messages as true more often than as lies (i.e., they are truth-biased), they get truths right more often than lies. The larger the truth-bias, the bigger the veracity effect, and the more truth accuracy and lie accuracy diverge.

Let's also make distinctions between cues, demeanor, and content. *Cues* are specific behaviors that are linked in some why to lying. Cues can be nonverbal like the amount of eye contact or the tone of voice, or they can be linguistic like first person pronouns. Cues might actually distinguish between truths and lies, or they might be specific behaviors people use to form opinions about honesty-deception or they might be part of folk wisdom about how liars behave. In any case, the link between cues and honesty-deception (be it actual or perceived) is probabilistic, not deterministic. For example, one potential cue is the number of details provided in a statement. On average and all things being equal, research shows that honest statements tend to include more details than lies (DePaulo et al., 2003). But, this does not mean that all statements lacking details are lies. If I am honestly describing one of my own studies I will probably provide more details than if I am describing someone else's research that I just skimmed. But, all other things being equal, statements with more details are more likely to be honest.

While cues refer to specific behaviors (like eye contact or the number of details), *demeanor* refers to a more general impression of a communicator. Demeanor is based on constellations of cues. People with an honest demeanor come off as competent, composed, friendly, and forthright. People with dishonest demeanors seem shifty, nervous, guilty, and lacking in confidence and commitment to what they are saying. What my research shows is that individual cues do not travel alone, and that when it comes to perceptions, demeanor (sets of related cues) have a powerful impact on how honest we think someone is (Levine et al., 2011).

Both cues and demeanor can be contrasted with *communication content*. While demeanor is about how something is said, content involves what is said. So, in the first paragraph when I challenged you to read the chapter to see if the title was accurate, reading the chapter is an issue of content. Details provide a good example of the distinction between cues and content. The number of details is a cue. As a cue, it does not matter what those details are. We just count the number of details. What the details are is content. For example: I am drinking my morning coffee as I type this paragraph. My coffee is pale green, quite salty, and frozen solid. See the three details in my description of my coffee? If I say "my coffee is now room temperature" there is just one detail. In this example, going by cues (more details = more likely honest) leads to a different conclusion than going by content. Coffee is neither green nor salty (at least I don't drink green, salty coffee), and liquids that are frozen solid are especially not drinkable. Coffee being room temperature,

in contrast, is quite plausible. This distinction between content and cues is very important in understanding deception detection.

Finally, let's talk about *diagnostic utility*. Diagnostic utility refers to the effectiveness of something or approach in deception detection accuracy. For example, I previously mentioned that averting eye gaze appears to have no utility (DePaulo et al., 2003). If gaze was the only thing you used to detect deception, your accuracy would be chance. The number of details has some utility. The correlation between honesty and the number of details is statistically significant and in the small to moderate range ($r = .15$, DePaulo et al., 2003). Obviously, if you want to detect lies accurately, you will want to base your judgments on things that have the most diagnostic utility.

Theories of Deception Detection

Theories of human deception detection can be sorted into three categories. Let's call these three sets (a) cue theories (b) the self-presentation approach, and my own (c) Truth-Default Theory.

Cue theories were historically first, and versions of cue theories are the most widely accepted (but also most controversial) theories. Cue theories have far and away produced the most research. Cue theories also align closely with folk belief about deception. The logic of cue theories holds that honest communication is psychologically different from deceptive communication. The psychological differences between truth telling and lying (e.g., emotions, cognitive effort, arousal level, type of strategic thinking) give rise to tell-tale behavioral signals (i.e., cues). Therefore, deception can be detected, albeit probabilistically and indirectly, by observation of cues.

For example, the most common folk belief about deception cues is that liars won't look you in the eye (Bond & The Global Deception Research Team, 2006). Gaze aversion is a behavior linked with feelings of shame. People who feel shame don't look others in the eye. They instead avert their gaze. So, if liars felt ashamed, we would expect them not to have as much eye contact as an unashamed honest speaker.

The first cue theory was described by Ekman and Friesen (1969). Ekman's idea (see Ekman, 2009 for the latest iteration) is that liars have different emotional experiences than honest communicators. Liars feel guilty and fear detection. They might experience delight at duping others. These emotions are behaviorally signaled, for example, through very brief micro-facial expressions. Spotting emotional displays triggered by lying provides behavioral clues that a person might be dishonest.

In 1981, cue theory thinking was expanded substantially by Zuckerman et al. They argued for four key psychological differences that link, explain, and predict how deception is signaled by specific cues. Relative to honest communication, lies were more arousing, more emotional, involved more cognitive effort, and involved more efforts at strategically controlling one's own behavior. Cues linked with any of these four psychological differences could be used to spot lies.

Interpersonal Deception Theory (IDT, Buller & Burgoon, 1996) lumped the internal psychological process described by Zuckerman et al. into two categories; strategic and non-strategic cues. The non-strategic cues were diagnostic and gave liars away. Non-strategic cues were specific behaviors signaling things like anxiety, arousal, or increased cognitive effort. Strategic behaviors, in contrast, were things deceivers did to act honest thereby fooling conversational partners. Cues might only give away less skilled deceivers. Skilled deceivers pick up on other's suspicion, strategically adapt, and act even more honest than honest senders thereby fooling less skilled receivers.

Another big part of IDT is that the extent of interactivity between sender and receiver is important in deception detection. IDT holds that face-to-face communication is different than mediated communication not only because of access to cues (eye contact isn't relevant with text messages) but because there is more or less back-and-forth adaptation between the communicators in different media.

The most recent cue theory is Vrij's (2015) cognitive approach which focuses on lie detection in criminal investigations. According to Vrij, the most important internal psychological process is cognitive load. Lying requires more cognitive effort than honest communication. However, cues are too weak on their own and need to be enhanced. By adding cognitive load, or asking for more information or asking unanticipated questions, cues can be magnified and made more diagnostic.

DePaulo's (1992) self-presentation approach offers logic opposite to that of various cue theories. In cue theories, honest and deceptive communication is psychologically different. According to DePaulo, both honest and dishonest communicators are concerned with creating a desired favorable impression on others. That is, honest and deceptive communication is much more similar than different (a presumption shared by McCornack, Morrison, Paik, Wiser, & Zhu's, 2014 Information Manipulation Theory 2). Further, most people have sufficient social skills to pull off their desired self-presentation. Because honest and deceptive communication is psychologically similar and because most people have sufficient social skills, cues are not expected to be very diagnostic. In short, DePaulo's self-presentation approach made the case that cues should not provide a useful lie detection approach and people are consequently poor lie detectors.

My own Truth-Default Theory (TDT, 2014) provides yet another view. I argue that most people just passively believe most communication. The idea that someone might be lying often does not even come to mind unless there is a good reason to be suspicious. I further argue (and my research shows, e.g., Levine, 2010; Levine et al., 2011) that cues and demeanor are misleading and produce accuracy only a little bit better than chance. Rather than cues, my theory focuses on content and persuasion as paths to better lie detection. Also, truth-biased leads to the veracity effect. TDT holds that the proportion of truths and lies (i.e., truth–lie base-rates) matter. If people are exposed to mostly honest communication, accuracy is higher. Accuracy is lower as lies become more probable (Park & Levine, 2015; Levine et al., 2006).

Meta-Analysis

Now that we have important terms defined and an understanding of various theories, it is time to get into actual findings. Most of the rest of this chapter provides an overview of research results. But, before I do that, let me explain my approach.

In reviewing the literature, I rely heavily on findings from meta-analysis when possible. Meta-analysis is a method of averaging findings across many studies and looking for patterns of findings. When I am reading descriptions of findings, I often find sentences that say this study found such-and-such. I often find this frustrating because I often know of several other studies that didn't find that. It looks to me like some authors pick evidence selectively to make their point. Other authors provide all sides. But, that can be frustrating too because we don't know what to believe when some studies find one thing and others don't. The findings are mixed. What does that mean? When faced with sets of conflicting findings, it is hard to know what to believe.

By focusing on meta-analysis, I avoid cherry-picking supportive findings and providing a biased account. I also get to see the bigger picture and that avoids the unsatisfactory "mixed results" or "it's too complicated" conclusions. Because there are huge numbers of prior studies, meta-analysis lets us see the trends that hold up over time.

Pre-2006 Findings (The Conventional Accepted Wisdom)

Looking back, the year 2006 was a turning point in deception research. That was the year that a highly influential meta-analysis by Bond and DePaulo (2006) was published. Bond and DePaulo's findings were based on huge numbers of previous studies and the findings provided a neat and coherent account of results. When adjusting for sample size, pretty much all the studies prior to 2006 led to the same conclusion. Old arguments were settled by the evidence. The findings were clear. But, 2006 was also a turning point for another reason. Around that time, the thinking began to change and new and different findings began to appear. Just as Bond and DePaulo had seemed to settle things, the evidence began to shift again. Looking back, there now seems to be two sets of findings, pre-2006 and post-2006 leading to two different sets of conclusions. The two sets of conclusions seem to contradict each other, but I think they are better understood as applying to different approaches to deception detection.

If you read lots of published deception detection experiments, you will come across the claim that people are no better than chance at detecting deception. This is wrong! But, it is not far from accurate. What research finds is that people are indeed better than chance, but not by very much. But that conclusion depends on what we mean by "very much." What Bond and DePaulo reported was that across all the studies they examined (which included just under 300 data sets), the average accuracy was just under 54%.

Statistically, that 54% average is very different from a 50–50 chance at odds of less than a ten-thousand to one. The effect size for the difference is $d = .42$, a statistically reliable difference of moderate size. But, at face value, accuracy of 54% means getting 11 out of 20 right rather than the 10 out of 20 expected by chance. An improvement of 1 in 20 just does not seem very impressive. So, the big take-home finding from Bond and DePaulo was that people are just slightly better than chance at detecting deception.

Accuracy findings prior to 2006 were normally distributed around the across-study average of 54% with a standard deviation of 6. A majority of prior findings fell between 50% and 58%. Findings much above or below that range were unusual and mostly from very small-sample data.

Now a few qualifications. The 54% is total accuracy. But, people were truth-biased, judging 55% of the messages as honest. This leads to the veracity effect. Accuracy for just honest messages was 61% and lie accuracy was 47%. Note that when the two are averaged $(61 + 47 \div 2)$ we get the 54% for total accuracy. The 54% is truth–lie discrimination. People are worse when it comes to correctly spotting lies per se.

The other noteworthy aspect of the findings was just how robust slightly better than chance accuracy findings were. The findings were consistent across media types (video only = 50%, audio only = 54%, audiovisual = 54%, face-to-face = 53%), for motivated and unmotivated senders (53% either way), for spontaneous and prepared lies (53% vs. 54%), regardless of access to baseline behaviors (53% vs. 55%) and for student and expert judges (53% vs. 54%). Accuracy was chance level if the judges had just video with no audio, otherwise average accuracy ranged from 53 to 55% regardless of other considerations that you might think would matter. For example, you might think that accuracy would improve with expert judges, motivated liars, exposure to honest baseline communication for a comparison, or when senders could not preplan their lies. If you thought such things (and many researchers did), Bond and DePaulo's results were disconfirming. Slightly better than chance accuracy ruled the day (at least in experiments published prior to 2006). Research shows that slightly better than chance accuracy holds for text, audio-only, audio-visual mediated, and face-to-face communication, more and less motivated liars, spontaneous and planned lies, college students and experts, and trained and untrained judges.

Caveats and Ecologies

It is hard to overstate the impact of Bond and DePaulo's (2006) findings. They took a while to settle in, but they were compelling. Over time they sunk in and got repeated and re-repeated. The conclusion drawn was that people are poor lie detectors.

Let me clearly say that I don't doubt Bond and Depaulo's findings at all. But, I nevertheless think the conclusion drawn from the findings is wrong in a subtle way. What the findings show is that people are just slightly better

at distinguishing truths and lies (total accuracy) in the types of lie detection experiments done before 2006. Findings are always limited by the methods used to obtain results. What happened after 2006 is that some research started trying new approaches. Let me explain.

One of the things that most of the prior studies had in common is that judges were typically shown an equal number of truths and lies. That is, if you are a subject in a lie detection experiment, there is usually an exactly 50% chance that any given message will be a lie. But, how does that match with everyday communication. My research shows that people are mostly honest most of the time (Serota & Levine, 2015; Serota, Levine, & Boster, 2010). Everyday communication tilts much more toward honesty compared to communication in laboratory settings. This puts a different light on findings of truth-bias. Research participants' judgments align more closely with everyday communication than does the communication in the laboratory. If we change up the ratio of truths and lie to be more realistic, accuracy systematically improves (Levine et al., 2006). So, maybe accuracy isn't quite as poor as research suggests.

Another characteristic of most deception detection experiments is the use of instructed lies, determined by random assignment. In most deception detection experiments, senders are randomly assigned to lie or tell the truth and they do so because they are told to. This makes things like character and motivation to lie irrelevant. Yet people lie for a reason (Levine, Kim, & Hamel, 2010), usually other than being told to do so based on random assignment. More recent research suggests that when people have access to information about motives, accuracy improves (Bond et al., 2013; Levine, Kim, & Blair, 2010). Furthermore, if senders in deception research who are instructed to lie end up confessing, their data is thrown out for a failure to follow instruction. Outside the laboratory, if a liar is persuaded to honestly confess the truth, this would be a successful detection.

A final thing to mention about prior lie detection experiments is that people are presumed to detect lies based on cues and demeanor. This makes really good sense from the perspective of the cue theory that had guided most of the research. In the typical experiment, content is of little value. There is no chance for fact-checking. Senders cannot be persuaded to honesty by confessing their deceptive ways. Instead, deception detection is limited to immediate judgments based on cues and demeanor.

How People Really Detect Lies

In 2002, Hee Sun Park and her colleagues tried looking at deception detection in a different way. Instead of experimentally having judges watch a bunch of interviews and try and assess honesty or asking about folk wisdom regarding what liars do, she asked her research participants to recall a recent occasion when they had detected a lie. How did they find out they were lied to?

The results were stunning. First, very few of the discovered lies that people recalled were detected at the time based on cues or demeanor. Instead, most

lies were only uncovered after the fact, sometimes long afterward. And, the most common discovery methods involved either the use of external evidence (e.g., physical evidence, information from informants, prior knowledge) or the liar confessing their lie and telling the truth. That is, the vast majority of recalled instances of detected lies involved communication content, not cues.

Park, Levine, McCornack, Morrison, and Ferrara, (2002) argued that one reason people do poorly in deception detection experiments is because lie detection experiments exclude the type of lie detection strategies that would-be lie detectors actually use in non-research settings. According to her, the proper conclusion to draw from meta-analysis is that people are slightly better than chance in lie detection tasks requiring real-time assessment of honesty based exclusively on cues and demeanor. People may well be better lie detectors in a longer game where evidence can be used and where senders might come clean. It is also worth noting that Park et al.'s findings have been independently and cross-culturally replicated with student and non-student samples (Masip & Herrero, 2015).

New and Improved Accuracy

Since the publication of the 2006 meta-analysis, at least two dozen published experiments suggest that improved deception detection accuracy is possible (Levine, 2015). There now seems to be several paths to improved accuracy in human lie detection. It is not that the old findings were somehow wrong, but rather that older research focused exclusively on cue-based lie detection.

Among the first major breakthroughs was the *strategic use of evidence* technique (SUE; Granhag, Stromwal, & Hartwig, 2007). In the SUE technique, the judge has some relevant prior knowledge. That is, the judge knows some facts and evidence relevant to assessing the honesty of the communication in question. But, the judge does not tell the sender what they know. The judge questions the sender. If the sender contradicts the known facts, the judge can toss out the facts bit by bit to see if the sender's story changes. Honest senders, of course, say things that align with the known (but undisclosed) facts. Statements that contradict the known facts are suspected and liars tend to shift their stories as the facts are gradually revealed. SUE has been used to obtain accuracy over 80% (Granhag et al., 2007).

Research also shows that content can be effective even when direct evidence is lacking. But, to make effective use of communication content, people need proper *context* (Blair, Levine, & Shaw, 2010) or familiarity with the situation (Levine & McCornack, 2001; Reinhard, Sporer, & Scharmach, 2013; Reinhard, Sporer, Scharmach, & Marksteiner, 2011). The idea is that statements taken out of context are more susceptible to misinterpretation, while taking what is said in context makes the content more informative. Being familiar with the situation facilitates understanding communication content having such context and situational familiarity allows us to meaningfully assess plausibility. Does the message make sense given what we know

about what is typical? Does what is said sound too good to be true? Does it stretch credulity? Blair et al. (2010) showed adding even a little important contextual information can increase accuracy in deception detection experiments to around 75%. Levine, Blair, and Clare (2014) further show that context-informed questioning can also produce accuracies in the 71–79% range.

Another aspect of content in context and situational familiarity that appears important is information about a sender's motives. Most people don't lie when their goals align with the truth; instead, people are more likely to lie when the truth is a problem (Levine et al., 2010). So, if we know something about the situation, we might be able to ascertain the extent to which someone has a motive to lie. Because having a motive to lie greatly increases the odds that a person so motivated will lie, knowledge of potential motive should greatly improve the odds of detecting the lie. Recent research strongly supports this idea, reporting accuracies of 86–97% when people have knowledge of motive (Bond, Howard, Hutchison, & Masip, 2013; Levine et al., 2010).

Finally, the ability to persuade senders to honestly confess their lies and tell the truth can greatly improve lie detection. There is little direct experimental evidence because almost all deception detection experiments discard data from confessors. This said, the divergent results of two recent experiments (Dunbar et al., 2015; Levine et al., 2014) are suggestive. The two experiments used a very similar setup involving experts questioning suspected student cheaters. In the Dunbar et al. experiment data from cheaters who confessed were discarded. In the Levine et al. experiments, the experts were allowed to seek confessions as a strategy, and honest confessions that were believed were counted as correct and contributed to accuracy. The Dunbar et al. study reported accuracy of 59%. The accuracies obtained by Levine et al. (2014) in their two experiments were 97 and 100%. While there are differences between the experiments other than just persuasion as a viable method, confession-seeking as a method appears to be a strong contributor to the difference in results.

Conclusions

Now that you have made it to the end of the chapter, you know that this chapter was indeed about human deception detection. Hopefully, several popular misconceptions have been revealed and you now have a state of the art understanding of the social science of lie detection. You should now know about cues and content, and how these play into deception detection.

Prior findings suggest that accuracy in deception detection rests on the lie detection approach. Different approaches lead to different findings. Most deception detection research before 2006 was based on cue theories. The idea is that truth telling and lying are psychologically different. Liars have different emotions, or liars are more strategic, or lying requires more cognitive

effort. These different internal psychological processes are signaled by specific nonverbal or linguistic behaviors called cues. These cues, then, can be used to detect lies. Research investigating cue-based lie detection yields very consistent results. Cue-based lie detection is statistically better than chance, but poor in an absolute sense. The often cited 54% accuracy refers to cue-based lie detection.

Over the last decade, improved accuracy has been documented in at least two dozen published research articles. Strategic use of evidence, content in context or situational familiarity, and persuasion-based approaches have all produced accuracy levels over 70%. What these new approaches all have in common is a movement away from cues to focus on communication content.

Can people tell when someone is lying? Sometimes. Going by folk wisdom like a liar won't look you in the eye is no better than flipping a coin. Other types of real-time nonverbal and linguistic cues have some, albeit limited, usefulness. But, listening to what is said in context yields better accuracy. Having some hard evidence and being able to persuade a potential liar to be honest is even better, presuming the investigative skill to gather valid evidence and the communication skills to induce honesty.

References

Blair, J. P., Levine, T. R., & Shaw, A. J. (2010). Content in context improves deception detection accuracy. *Human Communication Research, 36,* 423–442.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234.

Bond, C. F., & The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74.

Bond, C. F., Jr., Howard, A. R., Hutchison, J. L., & Masip, J. (2013). Overlooking the obvious: Incentives to lie. *Basic and Applied Social Psychology, 35,* 212–221.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111,* 203–243.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118.

Dunbar, N. E., Jensen, M. L., Burgoon, J. K., Kelly, K. M., Harrison, K. J., Adame, B. J., & Bernard, D. R. (2015). Effects of veracity, modality, and sanctioning on credibility assessment during mediated and unmediated interviews. *Communication Research, 42*(5), 649–674.

Ekman, P. (2009). *Telling lies.* New York, NY: W. W. Norton.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32,* 88–106.

Granhag, P. A., Stromwal, L. A., & Hartwig, M. (2007). The SUE technique: The way to interview to detect deception. *Forensic Update, 88,* 25–29.

Levine, T. R. (2010). A few transparent liars. In C. Salmon (Ed.), *Communication Yearbook 34* (pp. 41–61). Thousand Oaks, CA: Sage.

Levine, T. R. (2014). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392.

Levine, T. R. (2015). New and improved accuracy findings in deception detection research. *Current Opinion in Psychology, 6,* 1–5.

Levine, T. R., & McCornack, S. A. (2001). Behavioral adaption, confidence, and heuristic-based explanations of the probing effect. *Human Communication Research, 27,* 471–502.

Levine, T. R., Blair, J. P., & Clare, D. (2014). Diagnostic utility: Experimental demonstrations and replications of powerful question effects and smaller question by experience interactions in high stake deception detection. *Human Communication Research, 40,* 262–289.

Levine, T. R., Kim, R. K., & Blair, J. P. (2010). (In)accuracy at detecting true and false confessions and denials: An initial test of a projected motive model of veracity judgments. *Human Communication Research, 36,* 81–101.

Levine, T. R., Kim, R. K., & Hamel, L. M. (2010). People lie for a reason: An experimental test of the principle of veracity. *Communication Research Reports, 27,* 271–285.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs, 66,* 125–144.

Levine, T. R., Kim, R. K., Park, H. S., & Hughes, M. (2006). Deception detection accuracy is a predictable linear function of message veracity base-rate: A formal test of Park and Levine's probability model. *Communication Monographs, 73,* 243–260.

Levine, T. R., Clare, D., Blair, J. P., McCornack, S. A., Morrison, K., & Park, H. S. (2014). Expertise in deception detection involves actively prompting diagnostic information rather than passive behavioral observation. *Human Communication Research, 40,* 442–462.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., et al. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research, 37,* 377–403.

Masip, J., & Herrero, C. (2015). Police detection of deception: Beliefs about behavioral cues to deception are strong though contextual evidence is more useful. *Journal of Communication, 65,* 125–145.

McCornack, S. A., Morrison, K., Paik, J. E., Wiser, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology, 33,* 348–377.

Park, H. S., & Levine, T. R. (2015). Base-rates, deception detection, and deception theory: A reply to burgoon (2015). *Human Communication Research, 41,* 350–366.

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs, 69,* 144–157.

Reinhard, M., Sporer, S. L., & Scharmach, M. (2013). Perceived familiarity with a judgmental situation improves lie detection ability. *Swiss Journal of Psychology, 72,* 43–52.

Reinhard, M., Sporer, S. L., Scharmach, M., & Marksteiner, T. (2011). Listening, not watching: Situational familiarity and the ability to detect deception. *Journal of Personality and Social Psychology, 101,* 467–484.

Serota, K. B., & Levine, T. R. (2015). A few prolific liars: Variation in the prevalence of lying. *Journal of Language and Social Psychology, 33,* 431–440.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of reported deception. *Human Communication Research, 36,* 1–24.

Vrij, A. (2015). A cognitive approach to lie detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: Current challenges and new approaches* (pp. 205–229). Chichester, UK: Wiley.

A Review of Meta-Analyses About Deception Detection

R. Weylin Sternglanz, Wendy L. Morris, Marley Morrow and Joshua Braverman

If you have ever read a book or watched a television show depicting lie detection, or have had a casual conversation regarding the best way to figure out if someone is lying, you have likely heard many different theories, some of which conflict with each other. Even social psychologists who read journal articles on deception detection studies may find the literature confusing; a PsycInfo search on the term "deception detection" produces 1694 sources, including 1251 peer-reviewed journal articles. A systematic summary of results across the many studies on this topic may be the best way to make sense of the findings. *Meta-analyses* are one of the most effective methodological tools for summarizing and quantifying scientific effects across studies.

In a meta-analysis, the statistical findings from multiple studies are combined together in order to examine how robust an effect is across a variety of experimental paradigms. Researchers can examine not only the size of an effect, but

R. W. Sternglanz (✉)
Nova Southeastern University, Davie-Ft. Lauderdale, FL, USA
e-mail: sterngla@nova.edu

W. L. Morris
McDaniel College, Westminster, MD, USA
e-mail: wmorris@mcdaniel.edu

M. Morrow
Impact Analytics, Boca Raton, FL, USA

J. Braverman
University of Minnesota, Minneapolis, MN, USA

also whether the effect is moderated by participant traits (e.g., gender, age), experimental variables, and study settings. The main advantage of a meta-analysis over a literature review is that meta-analyses are less subjective and allow for more precise quantitative conclusions (Rosenthal & Rosnow, 1991). Meta-analyses are also generally considered superior to single empirical studies because of their increased statistical power and the fact that they usually do not rely on a particular researcher or experimental paradigm. In meta-analyses, effect sizes are typically presented using either a Pearson's r correlation coefficient (to assess the degree of relationship between two variables) or Cohen's d (to assess the difference between conditions, calculated as the difference between means divided by the pooled standard deviation). Roughly, an effect size of $d = .20$ (or $r = .10$ to $.30$) is considered small, an effect size of $d = .50$ (or $r = .30$ to $.50$) is moderate, and an effect size of $d > .80$ (or $r > .50$) is large (Cohen, 1988). These effect size measures allow for easy comparison across different types of studies and enable readers to have a sense of how strong or weak a relationship or effect is. Thus, meta-analyses are uniquely useful in providing precise estimates of an entire literature— far more useful than simply relying on whether or not a finding is statistically significant within a given study or set of studies (Lipsey & Wilson, 2001).

Indeed, one reason we wanted to write this chapter is to have a single source to provide to our students that can summarize social science's best answers to the "big" questions in deception detection research, such as: How accurate are people at detecting deception? Can experience, training, or circumstances make "perceivers" (people attempting to discriminate between truthful and deceptive communications) more accurate? What are the actual signs that a "sender" (the person who produces a truthful or deceptive communication) is lying, and what signs do perceivers *think* indicate that a sender is lying? Are polygraph machines, brain-imaging techniques, or other tools effective ways to enhance lie detection? Below, we summarize and interpret meta-analyses conducted to address these questions.

DECEPTION DETECTION ACCURACY AND MODERATORS OF ACCURACY

For laypeople interested in deception detection, perhaps no question is more important than knowing how likely they are to detect deception. Can perceivers discriminate between truthful and deceptive communications at substantially above-chance levels? If so, under what circumstances are perceivers more or less accurate? The meta-analyses below address these questions.

Meta-Analyses of Deception Detection Accuracy

Three of the earliest meta-analytic analyses on deception detection accuracy (DePaulo, Zuckerman, & Rosenthal, 1980; Kraut, 1980; Zuckerman, DePaulo, & Rosenthal, 1981) found results that have been replicated over the years. First, these analyses found that deception detection accuracy is only slightly better than chance. Second, the analyses showed that, contrary

to popular belief, voice and body cues were more useful than facial cues in detecting signs of deception (Zuckerman et al., 1981). In an unpublished doctoral dissertation, Kalbfleisch (1985) confirmed these findings and found early evidence of the tendency to see most communications as truthful, a finding subsequently labeled the *truth bias* (Levine, Park, & McCornack, 1999). Subsequent summaries of deception detection accuracy (e.g., Vrij, 2000) reported similar findings to these early analyses.

In 2006, Bond and DePaulo conducted a large-scale meta-analysis, including 206 studies and 24,483 perceivers of truthful vs. deceptive communications. This paper has been cited 1333 times as of April 2018, according to Google Scholar, and is generally considered the gold standard when it comes to measuring deception detection accuracy. Bond and DePaulo (2006) systematically gathered every known analysis (both published and unpublished) on perceivers' accuracy at discriminating between truthful and deceptive communications of strangers; studies in which judges received experimental training, instructions about how to detect deception, or special aids (such as a polygraph reading or behavioral codings) were excluded. There were 177 independent samples of senders and 384 independent samples of perceivers. Twelve percent of the perceivers had occupational expertise in detecting deception (i.e., about 2842 experts).

Across all 292 samples used in the meta-analysis, the mean accuracy in discriminating truthful from deceptive communications was approximately 54% (when 50% is chance), with an effect size of $d = .40$ when deceptiveness was measured on a continuum. The highest mean percentage correctly attained in any sample was 73% and the lowest was 31%. As found in earlier meta-analyses, perceivers demonstrated a truth bias; they correctly classified 61.3% of truthful messages as truthful, but only 47.6% of deceptive messages as deceptive. Thus, accuracy rates in any given study may depend on the number of truthful versus deceptive statements made by senders, with higher-accuracy rates when senders tell relatively few lies. Bond and DePaulo (2006) also confirmed that deception detection accuracy was lower when judgments were made via video rather than via an audiovisual medium ($d = -.44$), audio-only medium ($d = -.37$), or from written transcripts ($d = -.28$); accuracy did not differ significantly between transcript, audiovisual, or audio presentations. Accuracy rates may be lower when visual cues are provided because senders make conscious attempts to control the way they appear when lying; on the other hand, accuracy rates may be higher when audio cues are provided because audio cues may be more difficult for senders to control. Ironically, senders who were motivated to get away with their lies were actually slightly more likely to be detected than senders who were not motivated ($d = .17$), possibly because motivated senders display more signs of fear or nervousness while lying. Perceivers were more accurate in judging unplanned rather than planned messages ($d = -.14$). Additionally, planned messages appeared more truthful than unplanned messages ($d = .13$). Contrary to popular opinion, people with occupational expertise (e.g., law enforcement personnel,

psychiatrists, auditors) were not found to be superior to non-experts (e.g., college students) in discriminating lies from truths ($d=-.02$).

Are There Individual Differences in Judgments of Deception?

One of the earliest meta-analytic analyses of deception detection (Zuckerman et al., 1981) found no relationship between sex, self-monitoring, or Machiavellianism and perceivers' ability to discriminate truthful from deceptive communications. A subsequent meta-analysis by Aamodt and Custer (2006), including 206 studies and 16,537 participants, examined several other individual difference measures and found similar null results. Deception detection accuracy was not substantially related to confidence ($r=.05$), age ($r=-.03$), education ($r=.03$), or sex ($d=-.03$). Aamodt and Custer were particularly interested in the role of occupational expertise in detecting deception, but once again the results came up short. Law enforcement officers (including police, detectives, secret service agents, parole officers, and judges) were not significantly more accurate ($M=55.5\%$) than college students ($M=54.2\%$). Even among law enforcement personnel, years of experience did not predict deception detection ability ($r=-.08$).

One might think that adults are at least more accurate in detecting the lies of children. But surprisingly, a recent meta-analysis by Gongola, Scurich, and Quas (2017), which included 45 experiments with 7893 adult perceivers and 1858 child senders, found that adults detect only 54% of children's lies (i.e., no higher than the rate at which adults detect other adults' lies).

Aamodt and Custer's (2006) finding regarding confidence replicated a meta-analysis by DePaulo, Charlton, Cooper, Lindsay, and Muhlenbruck (1997), which extensively examined the relationship between deception detection judgments and confidence in those judgments. DePaulo et al. assessed 18 studies (including one unpublished manuscript) that reported correlations between continuous measures of confidence and accuracy. The finding across the 2972 perceivers, which included both college students and law enforcement personnel, was that the confidence–accuracy correlation did not significantly differ from zero ($r=.04$). In the six studies in which mean levels of confidence and accuracy could be compared, confidence was higher than accuracy. Thus, it appears people tend to be overconfident in their deception judgments, and their level of confidence says nothing about their accuracy. However, perceivers' confidence was related to other aspects of their deception judgments. Perceivers who were more confident in their judgments were more likely to perceive sender communications as truthful ($r=.17$). Perceivers' confidence in their judgments increased with the closeness of their relationship to the sender ($r=.19$), as predicted by theories that interpersonal perception in close relationships is based on an implicit sense of trust (McCornack & Parks, 1986). Men were significantly more confident about their deception judgments than women ($r=.15$) which is consistent with research showing that men are more confident—but not more accurate—than

women in a variety of interpersonal judgments (Patterson, Foster, & Bellmer, 2001). Finally, 8 studies demonstrated that perceivers were significantly more confident in their judgments when viewing truthful rather than deceptive communications ($r = .15$). This result supports the notion that lies can be detected indirectly (DePaulo & Morris, 2004; but see criticism of theories on unconscious lie detection, Street & Vadillo, 2016).

Might certain individuals differ in their ability to detect deception, even if these abilities are not aligned with such obvious traits as sex, age, personality measures, education, occupational expertise, or confidence? The notion that a small proportion of people are lie detection "wizards" is a tantalizing idea, supported by the research of O'Sullivan and Ekman (2004) and popularized in prime-time television shows such as *Lie to Me* (Baum, 2009). Bond and DePaulo addressed this notion in their 2008 meta-analysis of individual differences in deception detection ability. They developed sophisticated statistical techniques to determine whether the variation in perceivers' deception detection ability across studies was due to real differences in perceivers' ability or whether the variation was a result of random measurement error. Their analysis included 247 studies drawn from 89 published and 53 unpublished manuscripts. In total, the participants included 2945 senders and 19,801 perceivers. This large analysis indicated that the range in ability to detect deception was no greater than would be expected by chance. While some perceivers were much better (or much worse) than the standard 54%, lie detection accuracy was not a reliable individual difference. Of course, it is possible that a tiny fraction of lie detection wizards do exist, or that particular situations can produce high levels of accuracy without special training among a select few—but evidence for such claims has not yet been demonstrated meta-analytically.

Bond and DePaulo (2008) did find individual differences regarding aspects of deception judgments other than accuracy. For example, perceivers differ from each other in terms of how likely they are to label senders' communications as truthful (i.e., truth bias); the observed range for their judgments is 40% wider than would be expected by chance alone. In addition, senders differ in the ability to lie successfully. Most senders do not display obvious signs of deceptiveness, but there is a small proportion of senders who are extremely poor liars. The greatest individual difference among senders is their degree of credibility regardless of whether they are telling the truth or not; this range is 2.4 times as wide as what would be expected due to chance. Some people generally appear to be very honest or very deceptive, regardless of whether or not they are lying.

Can Deception Detection Accuracy Be Improved?

Readers who are seeking a "Pinocchio's nose," or surefire method to detect lies at near-100% rates, will be disappointed; while individual studies may claim to have found a technique to improve deception detection accuracy substantially without requiring perceivers to go through any special training,

most of the discrepancy across studies can be explained as random variation (Bond & DePaulo, 2006). However, Hartwig and Bond (2014) note that Bond and DePaulo's meta-analysis demonstrates the accuracy rate by human observers, rather than the "objective detectability" of lies. According to Hartwig and Bond, lies could theoretically be detected at a rate much higher than 54% if perceivers took multiple cues into account at once. Hartwig and Bond examined the degree to which lies could be detected if perceivers used all of the available behavioral cues. The researchers conducted a meta-analysis of 92 published and 33 unpublished studies (totaling 26,866 messages) that described a statistical prediction of deception from two or more visible, written, speech, vocal, or "global impression" cues. Lies could be objectively detected (using statistical algorithms and multiple cues) approximately 67% of the time on average, substantially higher than human perceivers' actual accuracy of 54%. However, this high detection rate only works in situations in which a large number of communications take place under similar circumstances, and is dependent on the ability to observe multiple (sometimes numerous) cues at once.

Might the context in which senders tell lies enable perceivers to obtain accuracy rates similar to that of the statistical algorithms? According to Hartwig and Bond (2014), this is unlikely; they found that lies are equally detectable regardless of senders' degree of motivation, whether senders are students or non-students, whether senders are communicating about feelings versus facts, or the setting in which the senders' communication takes place. Hartwig and Bond interpret this finding as evidence that deception detection accuracy rates are not an artifact of laboratory settings, because the detectability of lies remains consistent across a variety of settings and situational variables. In other words, the low-accuracy rate of human lie detection without special training is stable and generalizable.

What happens when perceivers *do* receive special training? This was the topic of a meta-analysis by Frank and Feeley (2003), who conducted an initial analysis of 20 studies (11 of which were published) on lie detection training. This meta-analysis compared 1072 participants who were trained in lie detection techniques during experiments to 1161 untrained perceivers. Frank and Feeley found that training did lead to a small gain in accuracy ($r = .20$). However, they note that there was considerable variance around this mean r value, with some studies showing much higher gains and some studies showing no gains in accuracy whatsoever due to training. Frank and Feeley suggested that future analyses should differentiate training that meets rigorous criteria from training that does not. Nine years later, this is exactly what Driskell (2012) did in his analysis of 16 published journal articles with 30 studies (total $N = 2847$). Using this updated dataset, Driskell found that training led to a moderate gain in accuracy ($d = .50$) but the accuracy was moderated by certain aspects of the training. First, training programs that included three components—instruction regarding signs of deception, practice in recognizing signs of deception, and feedback on perceivers' guesses

about senders' truthfulness—lead to high gains in accuracy ($d = .59$). Second, Driskell investigated the effects of the training content being taught to perceivers, looking specifically at actual cues to deception documented by DePaulo et al. (2003)—a meta-analysis discussed in more detail in the next section. Driskell found positive effects on accuracy when perceivers were correctly taught that senders are more likely to be lying if they exhibit tension or fidgeting, if their statements seem illogical, and if they make speech errors. Third, Driskell compared the effectiveness of training on perceivers with no special expertise (mostly college students) versus perceivers with experience in deception detection (mostly law enforcement personnel). Training was actually more effective for the perceivers without any experience. This may seem surprising, but Driskell points out that law enforcement training frequently focuses on stereotypical signs of deception, such as gaze aversion (Vrij, 2000), rather than empirically supported cues—which may actually lead law enforcement personnel to focus on some incorrect cues. Finally, Driskell found that training was more effective in teaching perceivers to detect lies about feelings and opinions than lies about transgressions. Most recently, Hauch, Sporer, Michael, and Meissner (2016) conducted an updated meta-analysis on training, based on 55 studies; unlike Driskell's analysis, Hauch et al. included unpublished findings and analyzed lie accuracy and truth accuracy separately. Hauch et al. found a small-to-moderate effect of training on detection accuracy of lies, but not truths. They also found that training was most effective when based on verbal content cues rather than nonverbal or paraverbal feedback. In sum, training that focuses on instruction regarding documented cues to deception, practice, and verbal content cues can be useful for detecting at least some lies. However, the degree to which training is effective in detecting lies in real-world interpersonal and forensic settings is less certain and a useful topic for future meta-analyses.

Actual Cues to Deception

Perhaps the most intriguing question about deception for researchers, law enforcement, and laypeople, beyond how to detect lies, is what the *real* cues to deception are. Are there reliable cues to deception, and if so, how strongly do they distinguish between truths and lies? The meta-analyses below address these issues, though the answers are less straightforward than the questions.

Nonverbal and Paraverbal Cues to Deception

The earliest meta-analyses on deception detection accuracy (e.g., Kraut, 1980; Zuckerman et al., 1981) documented some actual cues to deception, but we will focus on the updated and thorough meta-analysis of actual cues to deception conducted by DePaulo et al. (2003); this paper has been cited 2031 times as of April 2018, according to Google Scholar. This analysis included 120 independent samples (including 3 unpublished works), in which

1338 estimates of 158 verbal, paraverbal, and nonverbal cues to deception were assessed.

DePaulo et al. (2003) found dozens of cues that significantly differentiated between truthful and deceptive communications. In this review, we focus only on the most reliable findings—specifically, those that were statistically significant, based on at least 6 studies, and had an effect size d of at least .20. Compared to senders who told the truth, senders who lied exhibited more vocal tension ($d=.26$), spoke in a higher pitch ($d=.21$), and appeared more tense and nervous ($d=.27$). Vocal displays of tension and nervousness are among the most reliable "paraverbal" signs of possible deception. Paraverbal cues are vocal cues that accompany speech.

Sporer and Schwandt (2006) conducted a meta-analysis to analyze a small number of senders' paraverbal behaviors in great depth. Specifically, they examined message duration, number of words, speech rate, response latency, pauses, speech errors, speech repetitions, and vocal pitch across 41 manuscripts. Only two of these paraverbal cues were significantly related to deception overall: liars spoke in a higher pitch ($r=.10$) and took longer to begin responding to questions (i.e., greater response latency) ($r=.11$). Sporer and Schwandt also noted that the relationship of many cues to deception was heterogeneous—that is, they differed substantially depending on several moderators. For example, the aforementioned relationships of vocal pitch and response latency to deception were greater when senders talked at least in part about their feelings, rather than only facts. The relationship of paraverbal cues to deception also varied with the amount of senders' preparation and degree of motivation, as well as the type of experimental design.

In a similar meta-analysis, once again using data from 41 articles (54 studies), Sporer and Schwandt (2007) conducted an in-depth examination of the relationship between 11 nonverbal behaviors (blinking, eye contact, gaze aversion, head movements, nodding, smiling, self-touching, hand movements, illustrators, foot/leg movements, and postural shifts) and deception. Three of these behaviors occurred less often when people were lying than when they were telling the truth: nodding ($r=-.09$), hand movements ($r=-.19$), and foot/leg movements ($r=-.07$). Contrary to popular belief (Global Deception Research Team, 2006), averting one's gaze was unrelated to deception. As is the case with paraverbal cues, the effect sizes of the relationships between nonverbal cues and deception tend to be small and heterogeneous. Sporer and Schwandt found that the relationship between nonverbal cues and deception varied substantially with the content of the lie, whether or not senders were motivated, the degree to which senders prepared their statements, the type of experimental design, and the operationalization of the behaviors. For both paraverbal and nonverbal cues, context matters in their relationship to deception, and the correlations between these cues and deception are generally far smaller than most people expect.

Verbal and Content-Related Cues to Deception

Many of the cues that DePaulo et al. (2003) found differentiated truths from lies were not facial expressions, body movements, or tone of voice, but rather characteristics of the actual wording senders used, as well as general impressions perceivers had of senders. Compared to truthful senders, deceptive senders were perceived as displaying less verbal and vocal "immediacy," or signs of being clear and direct ($d=-.55$). Liars also seemed more uncertain ($d=.30$) and less emotionally involved in their statements ($d=-.21$). Liars made statements that seemed less plausible ($d=-.23$), less logical ($d=-.25$), and more internally discrepant or ambivalent ($d=.34$). DePaulo et al. interpreted these six findings as indicative that liars' communications are *less compelling* than those of truth-tellers. Liars also provided fewer details in their statements ($d=-.30$) and were perceived as making more negative statements and complaints ($d=.21$), leading perceivers to have a slightly more negative impression of liars than truth-tellers. Ironically, cues to deception are more obvious when senders are more motivated to succeed.

Computer-Identified Linguistic Cues to Deception

Can computers detect lies? This was the question asked by Hauch, Blandón-Gitlin, Masip, and Sporer (2015) in their meta-analysis of the linguistic cues to deception that can be detected by computer programs. Hauch et al. identified 79 cues from 44 studies (17 unpublished; total $N=3780$ senders) in which computer software programs (e.g., the Linguistic Inquiry and Word Count; Pennebaker, Booth, Boyd, & Francis, 2015) had been used to identify words indicative of deception. Hauch et al. found that liars used fewer words, as well as less-varied and complex words, supporting the notion that liars experience greater cognitive load (see Vrij, Fisher, & Blank, 2015). Liars also used more negative words (as well as more emotion words), which fits with DePaulo et al.'s (2003) finding that liars make more negative statements. Liars used fewer first-person pronouns and more second- and third-person pronouns, possibly indicating that liars are more likely to distance themselves from the events they discuss. Liars used fewer sensory and perceptual words, as well as fewer words related to their cognitive processes. As in other meta-analyses of actual cues to deception (DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007), effect sizes were generally small and heterogeneous. Effects were moderated by the type of event senders discussed, the degree of personal involvement, whether events discussed were positive or negative, the degree of interaction senders had with perceivers, and senders' level of motivation. As with nonverbal and paraverbal cues, context figures prominently in the relationship between linguistic cues and deception.

Do People Use Valid Cues to Detect Deception?

People generally perform only slightly better than chance at detecting lies, as demonstrated by the aforementioned Bond and DePaulo (2006) meta-analysis. There are two possible explanations for this finding. First, it is possible that people are unable to detect many lies because they rely on invalid cues to deception (perceived cues). Second, it is possible that people rely on valid cues to deception, but the dearth of valid behavioral cues and the small effect sizes associated with those cues lead to poor accuracy. Hartwig and Bond (2011) conducted a series of meta-analyses to evaluate which of these explanations receive greater empirical support.

Hartwig and Bond (2011) assessed the relationship between perceived cues to deception and actual cues to deception, examining 66 cues across 153 samples. The overall correlation between perceived and actual cues was $r = .59$, a moderate to strong relationship. When Hartwig and Bond limited their meta-analysis to "within-study evidence"—i.e., studies in which perceived cues and actual cues were measured within the same sets of perceivers and senders—the correlation between perceived and actual cues rose to $r = .72$, a very strong relationship. Hartwig and Bond found that deception detection accuracy was much more constrained by the lack of valid cues than by perceivers' tendency to use incorrect cues. In other words, perceivers mostly use the right cues to detect deception; limited lie detection accuracy can be attributed mostly to the fact that valid cues to deception are not very reliable.

Interestingly, Hartwig and Bond (2011) also found that the cues perceivers rely on when making deception judgments differ from the cues perceivers *claim* to rely on when making deception judgments. For example, perceivers frequently say that they use lack of eye contact to determine that a sender is lying; however, in actuality, lack of eye contact is only weakly related ($r = -.15$) to perceivers' judgments of deceptiveness. Consistent with classic findings that people are often misguided when reporting on their internal (often unconscious) cognitive processes (Nisbett & Wilson, 1977), people don't seem to know what cues they use when making deception judgments.

INTERROGATION TECHNIQUES USED BY LAW ENFORCEMENT

Law enforcement would be a much easier job if there were a highly accurate method to discern whether a suspect is lying or telling the truth. In the following section, we will describe the deception detection techniques used by law enforcement and provide meta-analytic data regarding the accuracy of each technique.

Polygraph—Control Question Test

The polygraph is often referred to as a "lie detector," implying that it can distinguish between truths and lies with a high degree of accuracy. More specifically, the polygraph measures certain physiological responses such as

respiration, pulse, blood pressure, and the skin's electrodermal response. The polygraph can be used with different types of questioning techniques. One such technique is the control question test (CQT) which is commonly used by law enforcement and government agencies in the US. In the CQT, a suspect's physiological responses during questions relevant to the crime are compared to their physiological responses during control questions that are unrelated to the crime. If the two patterns of physiological responses are significantly different from each other, the polygraph examiner is likely to conclude that the suspect is lying.

Kircher, Horowitz, and Raskin (1988) conducted a meta-analysis of the accuracy of the CQT which included 14 mock crime studies ($N=765$), including 2 unpublished studies. They focused their meta-analysis on mock crime studies rather than field studies because mock crime studies allow the researchers to know with certainty which participants are lying and which are telling the truth. Participants in mock crime studies are randomly assigned to commit a mock crime or are given information about a mock crime which they did not commit. In these studies, there is often an incentive, in the form of money or the avoidance of punishment, motivating both the guilty and innocent people to appear convincingly truthful in their claims of innocence. The meta-analysis of mock crime studies found that the overall detection accuracy of the CQT was 66%.

Because accuracy rates for the CQT in mock crime studies vary widely (21–87%), the purpose of the meta-analysis by Kircher et al. (1988) was to test whether the variability in accuracy is due to how ecologically valid the mock crime studies are—that is, how similar the conditions of mock crime studies are to conditions in the field. The meta-analysis found that the CQT is more accurate when the participants have some incentives or motivation to appear truthful ($r=.73$). The CQT was also more accurate when the samples in the mock crime studies were not predominantly college students but instead included members of the community, ex-offenders, and prison inmates ($r=.61$). Kircher et al. argue that the CQT may be less accurate with college students because college students may care less about the monetary incentives than would a prisoner or ex-offender. Finally, the CQT is more accurate when the polygraph examiners make decisions based on standard criteria used in the field such as using numerical coding and at least 3 charts of physiological data ($r=.67$).

Polygraph—Guilty Knowledge Test

Another questioning technique used with the polygraph is the guilty knowledge test (GKT), most frequently used in Japan and Israel (Ben-Shakhar & Elaad, 2003). The GKT is a series of multiple choice questions such as "what type of gun was used in the crime?" If the suspect's pattern of physiological responses is different when the correct answer is mentioned than when the incorrect answers are mentioned, this pattern would indicate that the suspect has personal knowledge of the details of the crime. The GKT can only be

used if the investigator knows what the answer to the question is (e.g., what type of gun was used) and if there is no conceivable way that an innocent person would have that information.

In a meta-analysis of 22 published studies ($N=1247$; unpublished studies were excluded from the analysis) conducted by MacLaren (2001), the overall accuracy rate of the GKT was 76%. When the meta-analysis was limited only to studies which included mock crimes, the accuracy rate increased to 82%. Similarly, in a meta-analysis of 80 studies conducted by Ben-Shakhar and Elaad (2003), the effectiveness of the GKT was higher in mock crime studies ($d=2.09$) than it was in the overall meta-analysis of all studies ($d=1.55$). The GKT was significantly more accurate in studies in which the GKT was implemented under conditions the researchers considered optimal. Those conditions were that the participants had a higher motivation to succeed, the participants had to verbalize a "no" response to the unselected options, and there were at least 5 guilty knowledge questions asked ($d=3.12$).

The Strategic Use of Evidence Technique

Law enforcement officials can also use deception detection techniques which do not rely on using a polygraph. When interrogators are in possession of highly incriminating evidence, they can use this information to their advantage when trying to detect deception. The Strategic Use of Evidence (SUE) technique involves not informing suspects that the interrogators are aware of the incriminating evidence until after the suspect has already provided his or her own version of events (Hartwig, Granhag, & Luke, 2014). The premise behind the SUE technique is that guilty people will avoid mentioning any information that could possibly be incriminating, whereas an innocent person will willingly share all information of which they are aware. For example, if there were a robbery at a store in a mall and suspects are asked to describe their recollection of that day, an innocent suspect would be more likely than a guilty person to mention shopping at the mall on the day of the crime. To a guilty person, this incriminating piece of information would be considered too aversive to mention and should be concealed.

When interrogators implement the SUE technique, they begin by asking open-ended questions such as, "Where were you on February 11th?" If the interrogator has evidence from a security camera that the suspect stopped at the mall, a failure to mention that detail would be considered a possible indication of deception. An interrogator using the SUE technique would look for two signs of deception: failure to mention incriminating information during the first telling of the story and inconsistencies between the suspect's statement and the known evidence. The SUE technique is most effective when the incriminating information is revealed to the suspect later in the interrogation process (Hartwig et al., 2014). If suspects already know that they were captured on a security camera, they can incorporate that detail into their

story in a way that doesn't make them look guilty. For example, they could say that they went shopping at the mall that day but fail to mention the particular store they went to. Withholding the incriminating evidence until late in the interrogation is likely to lead to more avoidance of the incriminating information and more inconsistencies between their statement and the evidence. Hartwig et al. (2014) conducted a meta-analysis ($N=599$) comparing the effectiveness of the SUE technique to a non-SUE technique using 8 mock crime studies (including 1 unpublished study). For both the SUE and non-SUE techniques, the statements of guilty people were more inconsistent with the evidence than the statements of innocent people, but the effect size was much larger when the SUE technique was used (non-SUE technique $d=1.06$; SUE technique $d=1.89$).

Although the SUE technique is very effective, its use is limited because it can only be used when the interrogator has incriminating evidence and the suspect is unaware that the evidence is known to the interrogator. In such situations where both of these requirements are met, SUE may be the most effective deception detection technique currently available which does not require a polygraph.

Increasing Cognitive Load

Lying requires more cognitive resources than telling the truth because a liar must actively create untruths, whereas a truth-teller simply has to describe existing memories (Vrij et al., 2015). Cognitive approaches to lie detection are based on the assumption that nonverbal cues may distinguish liars from truth-tellers precisely because those cues appear when a liar is using a lot of cognitive resources, a state known as cognitive load. When implementing a cognitive approach to deception detection, the interrogator or experimenter intentionally does certain things to make the task even more demanding for liars. For example, an interrogator might increase cognitive load by asking the suspect to tell the story backward, to make unwavering eye contact with the interrogator while telling the story, or to tell the story while doing another task simultaneously. An interrogator can also increase the difficulty of the task by asking suspects to add more details to their stories; a truth-teller can do this easily because there are many possible details to share, but a liar would need to create those details on the spot. Another way to increase cognitive load is to ask the suspect unanticipated questions. Liars often plan their answers to anticipated questions in advance which reduces their cognitive load when giving those answers during the interrogation. Suspects will be under higher cognitive load if they are asked questions they did not anticipate having to answer.

Vrij et al. (2015) conducted a meta-analysis of 14 studies to compare the effectiveness of the cognitive approach to a standard lie detection approach in which cognitive load is not intentionally increased. Their meta-analysis

indicated the cognitive approach was more accurate than the standard approach in accurately detecting lies (67% vs. 47%, $d = .53$), accurately detecting truths (67% vs. 57%, $d = .24$), and overall accuracy (71% vs. 56%, $d = .42$). Because cognitive approaches increase the difficulty of the task, suspects "leak" twice as many verbal and nonverbal cues to deception than when a standard approach is used (Vrij, Fisher, Blank, Leal, & Mann, 2016).

Content-Based Techniques

Content-based techniques were designed to differentiate between truthful and deceptive statements by examining the specific content shared by suspects. Content-based techniques are based on the assumption that statements about personal experience will include more detail than statements not based on actual experience. For example, a truthful story may be more likely to contain details about the context of the event, conversations that occurred, and recollections of one's mental state. Furthermore, because speakers of fabricated stories are especially concerned with appearing truthful, they may be less likely to correct their stories spontaneously or admit to forgetting some aspect of the event than would a truthful speaker. Fabricating a story and trying to appear truthful through self-presentation both require cognitive resources which leave fewer resources available to add extensive details to the story. Two types of content-based techniques are criteria-based content analysis (CBCA) and reality monitoring (RM). CBCA was developed to distinguish between true and fabricated statements, and it is considered admissible evidence in a court of law in the US and Western Europe. Although RM was originally developed as a way to distinguish between true and false memories, the technique has also been used in lie detection. With both of these techniques, there is a list of criteria that are used to judge the statements. A recent meta-analysis of 56 studies tested the effectiveness of CBCA and RM in detecting deceptive statements (Oberlader et al., 2016). Overall, the accuracy rate for these techniques was 70% ($d = 1.03$) and there was no significant difference between the effectiveness of the two techniques.

Limitations of Studying the Effectiveness of Lie Detection Techniques in Lab-Based Settings

While the above meta-analyses are very useful in giving us an idea of how accurate each method is, we must be very cautious about assuming these accuracy rates will be the same outside of the laboratory and in the field settings where professionals use the techniques to solve crimes. Experiments testing the accuracy of lie detection techniques typically take place in highly controlled laboratory environments so that the experimenter can randomly assign participants to be guilty and innocent, which enables experimenters to know with certainty whether a particular deception detection technique

was accurate in each case. This level of certainty in testing the accuracy of a lie detection technique is not possible in the field because professional law enforcement officials usually cannot know definitively which suspects are innocent and which are guilty. In an effort to increase the external validity of laboratory studies, researchers have used mock crime experiments which attempt to mirror real-world conditions by including incentives to appear innocent or punishments if one is found guilty (Ben-Shakhar & Elaad, 2003; Hartwig et al., 2014). However, given that the incentives and punishments associated with a real criminal investigation are far greater, the accuracy rates of the lie detection techniques above should be considered estimates rather than definitively conclusive.

NEUROSCIENTIFIC TECHNIQUES FOR LIE DETECTION

Technological advances over the past 20 years have provided new tools for studying brain activity during deception (Christ, Van Essen, Watson, Brubaker, & McDermott, 2009). Unlike older techniques for studying brain activity (via scalp-recorded event-related potentials), positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) allow researchers to study the specific brain regions being activated while participants engage in various forms of deceptive communication (Kozel et al., 2005). In their meta-analysis on the use of fMRI and PET techniques in 12 studies, including a total of 173 activation foci, Christ et al. sought to determine: (a) which regions of the brain were activated during deception, and (b) which aspects of executive control (working memory, inhibitory control, or task switching) were most important during the act of deception. Christ et al. found that 13 brain regions were more active during deceptive than truthful communication, and 8 of these 13 brain regions are located in or near the prefrontal cortex (PFC). This finding supports the theory that executive control processes play an important role in deception, because the PFC has a strong role in executive functioning. Some of these regions (specifically, the right and left inferior frontal gyrus [IFG] and insula, as well as the anterior cingulate cortex [ACC]) contribute to executive control generally, and thus, it is difficult to know if these regions have a specific role in deception per se, rather than just a role in all executive control functions. However, the left dorsolateral PFC, the right anterior PFC, and right posterior parietal cortex were associated with both deception and working memory, but not other executive functions. This indicates that working memory, more than inhibitory control or task switching, may play a particularly important role in deception. Furthermore, the insula and nearby (parainsular) regions of the brain, which are known to play a role in visceral responses such as blood pressure and heart rate, were activated during deception. As discussed in the section on polygraph techniques, these functions frequently accompany deception; thus, it is unsurprising that these regions of the brain are active

during deception. Finally, two regions of the brain unrelated to executive control were activated during deception—specifically, the left and right inferior parietal lobules. These regions of the brain have previously been implicated in selective attention and detection of important low-frequency events. These areas of the brain may play a role in maintaining attention in order to detect contexts in which deception is required.

There are various kinds of lies, involving various types of cognitive and emotional processes (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996); thus, it is likely that different regions of the brain are utilized for these different types of lies. Lisofsky, Kazzer, Heekeren, and Prehn (2014) conducted a meta-analysis of neuroimaging studies using PET and fMRI in an attempt to differentiate the regions of the brain utilized during socially interactive vs. non-interactive lies. Socially interactive lies, which the authors consider more ecologically valid, include tasks such as deceiving an interrogator about autobiographical information, making false promises to behave cooperatively in a trust/prisoner's dilemma game, and concealing knowledge about memories or knowledge when asked. Non-interactive lies include tasks such as lying about whether an object or word is recognized, or whether an everyday act has been performed correctly. Twenty-four studies, including 26 contrasts between truthful and deceptive statements ($N=416$), were included in the meta-analytic comparison between socially interactive vs. non-interactive studies. Studies were classified as having "social interactive" (as opposed to "non-interactive") deception paradigms: if (a) an interaction partner who gets deceived was present or imagined by the participant, or (b) there was a cover story designed to simulate a real-world interpersonal deception. The analysis compared the neural activity for the 15 social interactive study paradigms vs. the 11 non-interactive study paradigms. Consistent with the authors' predictions, the regions of the brain involved in social interactive (vs. non-interactive) deception involved working memory and inhibitory control. One of the regions of the brain that was more active during social lies was the ACC, which is known to play a role in detecting or monitoring conflict (a situation that may occur when people lie in social situations). The posterior superior temporal and angular gyrus was also more active during social lies; this region of the brain has been associated with social cognition and moral decision-making, including theory-of-mind processes (i.e., inferring the mental states of others). It is likely that people engage in social deception designed to fool an interaction partner by making inferences about the partner's mental state. The role of a part of the brain associated with moral decision-making may indicate that telling a lie in a social setting is considered a moral transgression, whereas non-interactive lies may not elicit that same sense. In addition to replicating many of the patterns found in the meta-analysis by Christ et al. (2009) regarding brain activity indicative of deception, Lisofsky et al. found that regions of the brain responsible for theory of mind and moral reasoning are utilized more in social interactive than non-interactive deception.

Findings from the aforementioned neuroscience meta-analyses show a good deal of consistency—perhaps unsurprisingly, because they included many of the same studies in their analyses. Brain-imaging techniques for deception detection have been heavily marketed as a potential cutting-edge tool to be used in business negotiations, protection against terrorists, and criminal trials. Can fMRI or PET technology serve as a useful lie detector in forensic or negotiation settings? Farah, Hutchinson, Phelps, and Wagner (2014) conducted a meta-analysis of lie detection studies using fMRI, with a focus on the practical and ethical implications of using brain-imaging tools in these applied settings. Their sample, which included 23 studies comparing responses to deceptive and truthful statements, indexed 321 foci in the brain. Farah et al. delineate several reasons why this technology may not be reliable in applied settings. First, although the meta-analytic findings are consistent, there is considerable variability in findings from study to study; no single brain region was active during deception in all the studies, or even almost all of the studies. Consistent with this finding, Gamer (2011) conducted a meta-analysis of fMRI studies using two different deception detection paradigms (22 studies, $N=408$) and found that the brain regions most active during deception depended heavily on the type of paradigm used to elicit deception. Second, for all deception detection paradigms, it is extremely difficult to determine whether differences in brain activity between the "lie" and "truth" conditions are due to the degree of truthfulness or to some other subtle difference between conditions. For example, Farah et al. point out that the frequency of motor response tended to be greater during deceptive than truthful statements; therefore, differences in neural activity may actually reflect brain activity associated with differences in motor actions. Similarly, differences in neural activity may be due to the greater cognitive load imposed by deceptive versus truthful statements. This could lead to false positives when participants are under cognitive load for reasons other than telling a lie. Another problem for practical applications is that fMRI studies may be particularly vulnerable to countermeasures—possibly much more so than other physiological lie detection measures such as the polygraph. For example, in one study Farah et al. reviewed, if participants made imperceptible finger and toe movements during their truthful and deceptive statements, accuracy fell to chance.

Even when lie detection via fMRI is reasonably accurate, it may still have low specificity, meaning that it is not useful in spotting low-probability events because too many false positives will occur (Farah et al., 2014).

Yet another issue with applying fMRI research to criminal investigations is that almost all participants in laboratory studies were college students with no diagnosed psychopathologies. Some violent criminal offenders, on the other hand, can be diagnosed with traits related to psychopathy and antisocial personality disorder. These diagnoses have been linked to structural and functional differences in brain activity, calling into question whether fMRI findings would differ for these populations (Farah et al., 2014).

The idea of sophisticated brain imaging as a tool to detect deception is appealing to the public at large. Indeed, at least two companies (No Lie MRI and Cephos) have recently started to offer fMRI lie detection services in business, personal, criminal, and national security settings (Farah et al., 2014). However, use of fMRI for lie detection in these real-world settings is almost certainly premature, and, thus far, fMRI evidence of deception is not generally accepted as evidence in criminal or civil court cases.

STATISTICAL LIMITATIONS TO META-ANALYSES ON DECEPTION

As we discussed earlier, meta-analyses are an exceptionally useful tool for quantitatively summarizing an entire field of research (Rosenthal & Rosnow, 1991) and this is particularly helpful in the study of deception detection, given the breadth of literature and the wide variety of experimental paradigms used. Nevertheless, meta-analytic techniques are not without flaws—flaws which are largely the result of problems with the entire paradigm of significance-testing across multiple areas of science (Simonsohn, Nelson, & Simmons, 2014).

One of the earliest documented flaws is the *file drawer problem*, which is the phenomenon that statistically significant results are much more likely to be published than nonsignificant results (Rosenthal, 1979). If, for example, one study demonstrates that a given manipulation improves perceivers' ability to detect deception, while ten other studies find no such effect, it is possible that the study showing the significant effect will be published, while the ten studies with null findings will go unpublished (i.e., sit in a file drawer), leading readers to infer that the manipulation does indeed improve deception detection. Meta-analyses are potentially quite useful in exposing these false positives, but only if all studies—including unpublished studies—are included in the analysis. A meta-analysis that systematically eliminates null results gives a skewed picture of the literature. Where possible, we noted when the meta-analyses we discussed made an effort to attenuate the file-drawer problem by including unpublished findings.

In addition to the file-drawer problem, other common researcher practices have been found to add to the rate of false positives. Many of these questionable practices were illustrated dramatically in a paper by Simmons, Nelson, and Simonsohn (2011) who demonstrated that flexibility in the collection, analysis, and reporting of data leads to a skewed publication record. When individual studies are biased in favor of a certain outcome, the meta-analysis of those studies will be biased as well (McShane, Böckenholt, & Hansen, 2016). Researchers have developed statistical techniques specifically to correct for these biases when conducting meta-analyses (e.g., Hedges, 1992; the "*p*-curve analysis" by Simonsohn et al., 2014). However, other researchers have found that these efforts to correct for publication biases in meta-analyses are not always adequate (Inzlicht, Gervais, & Berkman, 2015). In addition,

even when meta-analyzing the exact same literature, different researchers may use slightly different methods for choosing precisely how to combine effect sizes from multiple studies, leading to slightly different outcomes (e.g., see DePaulo et al., 2003; Sporer & Schwandt, 2006). Nevertheless, despite these flaws, most researchers agree that meta-analyses are still far more useful than single studies in ascertaining average effect sizes, the degree of heterogeneity within findings, and the role of moderators in study results (McShane et al., 2016).

Potential Future Meta-Analyses on Deception

There are at least three research topics in the deception literature that have not been meta-analyzed yet, but we hope they will be in the near future. First, to the best of our knowledge, there are no meta-analyses on the frequency with which people lie, despite some intriguing studies in this area (e.g., DePaulo et al., 1996). Second, we could not find meta-analyses on the role of relationship closeness in deception detection accuracy, although there are a number of published manuscripts in this area (e.g., Anderson, DePaulo, & Ansfield, 2002; Boon & McLeod, 2001; Levine & McCornack, 1992; McCornack & Levine, 1990; Morris et al., 2016; Sternglanz & DePaulo, 2004; also see Sternglanz & Morris, 2014, for a brief review of deception in friendships). Third, the efficacy of implicit or indirect deception detection is a hotly debated topic (DePaulo & Morris, 2004; Levine & Bond, 2014), and a meta-analysis may provide clarity on this issue. Finally, there are numerous topics related to deception in psychology (e.g., embellished resumes, infidelity, academic dishonesty, children's understanding of false belief tasks) and behavioral economics (e.g., game theory paradigms such as prisoner's dilemma) with broad literatures, some of which have been meta-analyzed. It would elucidate our understanding of deception to integrate findings from these analyses with meta-analytic findings on deception detection.

Conclusions

Meta-analyses have been conducted on a wide variety of topics related to deception detection; see Table 16.1 for a brief quantitative summary. Findings from these analyses indicate that people are generally only slightly better than chance at detecting deception, regardless of their personality traits, career experience, or confidence in their judgments. There are cues that probabilistically indicate when people may be lying, but only a small minority of liars display obvious cues. Limited deception detection accuracy can be attributed mostly to the fact that valid cues to deception are not highly reliable. Nevertheless, training programs that focus on documented cues to deception, verbal content cues, and practice can improve perceivers' ability to detect at least some lies. Additionally, computer programs and statistical algorithms

Table 16.1 Deception detection accuracy for meta-analyzed techniques

Deception detection technique	Accuracy	Primary meta-analyses
No special technique or training	54%, $d = .40$ No differences due to confidence, education, sex, personality, or occupational expertise	Bond and DePaulo (2006), Aamodt and Custer (2006), and DePaulo et al. (1997)
Statistical algorithms or computer programs (non-human)	67% under specific parameters	Hartwig and Bond (2014) and Hauch et al. (2015)
Empirically supported training in deception detection	$d = .33$ to $.59$	Driskell (2012) and Hauch et al. (2016)
Polygraph—Control Question Test	66%	Kircher et al. (1988)
Polygraph—Guilty Knowledge Test	76% in general, $d = 1.55$; 82% in mock crime studies, $d = 2.09$	MacLaren (2001) and Ben-Shakhar and Elaad (2003)
Strategic use of evidence	$d = 1.89$	Hartwig et al. (2014)
Cognitive approach	71%, $d = .42$	Vrij et al. (2015)
Content-based approach	70%, $d = 1.05$	Oberlader et al. (2016)
Neuroscientific techniques (fMRI, PET) Several regions of the prefrontal cortex, and especially brain regions associated with working memory, are consistently associated with deception.	Highly variable (near 100% with ideal experimental controls, but as low as chance if conditions are not ideal or the sender uses countermeasures)	Christ et al. (2009), Lisofsky et al. (2014) and Farah et al. (2014)

can detect lies better than human perceivers under specified conditions. Law enforcement tools such as polygraphs, strategic use of evidence, and increasing senders' cognitive load also improve lie detection, at least in controlled experimental settings. Neuroscientific techniques such as fMRI have demonstrated that areas of the brain associated with working memory are more active during deception; however, despite high levels of accuracy under specific highly controlled conditions, brain imaging is, at present, an unreliable tool for detecting real-world lies. In spite of some limitations, meta-analyses have been highly useful in summarizing the effect sizes, degree of heterogeneity, and moderators for the scientific study of deception detection.

References

* References marked with an asterisk indicate meta-analyses reviewed in this chapter.

*Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15,* 6–11.

Anderson, D. E., DePaulo, B. M., & Ansfield, M. E. (2002). The development of deception detection skill: A longitudinal analysis of same-sex friends. *Personality and Social Psychology Bulletin, 28,* 536–545. https://doi.org/10.1177/0146167202287010.

Baum, S. (Creator). (2009). *Lie to me* [Television series]. Los Angeles, CA: Twentieth Century Fox.

*Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the guilty knowledge test: A meta-analytic review. *Journal of Applied Psychology, 88,* 131–151. https://doi.org/10.1037/0021-9010.88.1.131.

*Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

*Bond, C. F., Jr., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477–492. https://doi.org/10.1037/0033-2909.134.4.477.

Boon, S. D., & McLeod, B. A. (2001). Deception in romantic relationships: Subjective estimates of success at deceiving and attitudes toward deception. *Journal of Social and Personal Relationships, 18,* 463–476. https://doi.org/10.1177/0265407501184002.

*Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. E. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex, 19,* 1557–1566. https://doi.org/10.1093/cercor/bhn189.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

*DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review, 1,* 346–357. https://doi.org/10.1207/s15327957pspr0104_5.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995. https://doi.org/10.1037/0022-3514.70.5.979.

*DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

DePaulo, B. M., & Morris, W. L. (2004). Discerning lies from truths: Behavioural cues to deception and the indirect pathway of intuition. In P. A. Granhag & L. A. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 15–40). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511490071.002.

*DePaulo, B. M., Zuckerman, M., & Rosenthal, R. (1980). Humans as lie detectors. *Journal of Communication, 30,* 129–139. https://doi.org/10.1111/j.1460-2466.1980.tb01975.x.

*Driskell, D. (2012). Effectiveness of deception detection training: A meta-analysis. *Psychology, Crime, & Law, 18,* 713–731. https://doi.org/10.1080/1068316X.2010.535820.

*Farah, M. J., Hutchinson, J. B., Phelps, E. A., & Wagner, A. D. (2014). Functional MRI-based lie detection: Scientific and societal challenges. *Nature Reviews Neuroscience, 15,* 123–131. https://doi.org/10.1038/nrn3665.

*Frank, M. G., & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research, 31,* 58–75. http://dx.doi.org/10.1080/00909880305377.

*Gamer, M. (2011). Detecting of deception and concealed information using neuroimaging techniques. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the concealed information test* (pp. 90–113). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511975196.006.

Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74. https://doi.org/10.1177/0022022105282295.

*Gongola, J., Scurich, N., & Quas, J. A. (2017). Detecting deception in children: A meta-analysis. *Law and Human Behavior, 41,* 44–54. https://doi.org/10.1037/lhb0000211.

*Hartwig, M., & Bond, C. F., Jr. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 4,* 643–659. https://doi.org/10.1037/a0023589.

*Hartwig, M., & Bond, C. F., Jr. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology, 28,* 661–676. https://doi.org/10.1002/acp.3052.

*Hartwig, M., Granhag, P. A., & Luke, T. (2014). Strategic use of evidence during investigative interviews: The state of the science. In D. C. Raskin, C. R. Honts, & J. C. Kircher (Eds.), *Credibility assessment: Scientific research and applications* (pp. 1–36). Oxford, UK: Academic Press.

*Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review, 19,* 307–342. https://doi.org/10.1177/1088868314556539.

*Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2016). Does training improve the detection of deception? A meta-analysis. *Communication Research, 43,* 283–343. https://doi.org/10.1177/0093650214534974.

Hedges, L. V. (1992). Modeling publication selection effects in meta-analysis. *Statistical Science, 7,* 246–255.

Inzlicht, M., Gervais, W. M., & Berkman, E. T. (2015). *Bias-correction techniques alone cannot determine whether ego depletion is different from zero: Commentary on Carter, Kofler, Forster, & McCullough, 2015.* Available at SSRN: http://ssrn.com/abstract=2659409 or http://dx.doi.org/10.2139/ssrn.2659409.

Kalbfleisch, P. J. (1985). *Accuracy in deception detection: A quantitative review* (Unpublished doctoral dissertation). Michigan State University, East Lansing, MI.

*Kircher, J. C., Horowitz, S. W., & Raskin, D. C. (1988). Meta-analysis of mock crime studies of the control question polygraph technique. *Law and Human Behavior, 12,* 79–90. https://doi.org/10.1007/BF01064275.

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry, 58,* 605–613. https://doi.org/10.1016/j.biopsych.2005.07.040.

*Kraut, R. (1980). Humans as lie detectors: Some second thoughts. *Journal of Communication, 30,* 209–218. https://doi.org/10.1111/j.1460-2466.1980.tb02030.x.

Levine, T. R., & Bond, C. F., Jr. (2014). Direct and indirect measures of lie detection tell the same story: A reply to ten Brinke, Stimson, and Carney (2014). *Psychological Science, 25,* 1960–1961. https://doi.org/10.1177/0956797614536740.

Levine, T. R., & McCornack, S. A. (1992). Linking love and lies: A formal test of the McCornack and Parks model of deception detection. *Journal of Social and Personal Relationships, 9,* 143–154. https://doi.org/10.1177/0265407592091008.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs, 66,* 125–144. https://doi.org/10.1080/03637759909376468.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

*Lisofsky, N., Kazzer, P., Heekeren, H. R., & Prehn, K. (2014). Investigating socio-cognitive processes in deception: A quantitative meta-analysis of neuroimaging studies. *Neuropsychologia, 61,* 113–122. https://doi.org/10.1016/j.neuropsychologia.2014.06.001.

*MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *Journal of Applied Psychology, 86,* 674–683. https://doi.org/10.1037//0021-9010.86.4.674.

McCornack, S. A., & Levine, T. R. (1990). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communication Monographs, 57,* 219–230. https://doi.org/10.1080/03637759009376197.

McCornack, S. A., & Parks, M. R. (1986). Deception detection and relationship development: The other side of trust. In M. L. McLaughlin (Ed.), *Communication yearbook 9.* Newbury Park, CA: Sage. https://doi.org/10.1080/23808985.1986.11678616.

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science, 11,* 730–749. https://doi.org/10.1177/1745691616662243.

Morris, W. L., Sternglanz, R. W., Ansfield, M. E., Anderson, D. E., Snyder, J. H., & DePaulo, B. M. (2016). A longitudinal study of the development of emotional deception detection within new same-sex friendships. *Personality and Social Psychology Bulletin, 42,* 204–218. https://doi.org/10.1177/0146167215619876.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259. https://doi.org/10.1037/0033-295X.84.3.231.

*Oberlader, V. A., Naefgen, C., Koppehele-Gossel, J., Quinten, L., Banse, R., & Schmidt, A. F. (2016). Validity of content-based techniques to distinguish true and fabricated statements: A meta-analysis. *Law and Human Behavior, 40,* 440–457. https://doi.org/10.1037/lhb0000193.

O'Sullivan, M., & Ekman, P. (2004). The wizards of deception detection. In P. A. Granhag & L. A. Stromwall (Eds.), *Deception detection in forensic contexts* (pp. 269–286). Cambridge, UK: Cambridge Press. https://doi.org/10.1017/cbo9780511490071.

Patterson, M. L., Foster, J. L., & Bellmer, C. (2001). Another look at accuracy and confidence in social judgments. *Journal of Nonverbal Behavior, 25,* 207–219. https://doi.org/10.1023/A:1010675210696.

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic inquiry and word count: LIWC 2015.* Austin, TX: Pennebaker Conglomerates, Inc. (www.liwc.net).

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641. https://doi.org/10.1037/0033-2909.86.3.638.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22,* 1359–1366. https://doi.org/10.1177/0956797611417632.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-Curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science, 9,* 666–681. https://doi.org/10.1177/1745691614553988.

*Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20,* 421–446. https://doi.org/10.1002/acp.1190.

*Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13,* 1–34. https://doi.org/10.1037/1076-8971.13.1.1.

Sternglanz, R. W., & DePaulo, B. M. (2004). Reading nonverbal cues to emotions: The advantages and liabilities of relationship closeness. *Journal of Nonverbal Behavior, 28,* 245–266. https://doi.org/10.1007/s10919-004-4158-7.

Sternglanz, R. W., & Morris, W. L. (2014). Deception in relationships: Friends. In T. R. Levine (Ed.), *Encyclopedia of Deception* (Vol. 2, pp. 806–809). Thousand Oaks, CA: Sage.

Street, C. N. H., & Vadillo, M. A. (2016). Can the unconscious boost lie-detection accuracy? *Current Directions in Psychological Science, 25,* 246–250. https://doi.org/10.1177/0963721416656348.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice.* New York: Wiley.

*Vrij, A., Fisher, R. P., & Blank, H. (2015). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22,* 1–21. https://doi.org/10.1111/lcrp.12088.

Vrij, A., Fisher, R., Blank, H., Leal, S., & Mann, S. (2016). A cognitive approach to elicit nonverbal and verbal cues of deceit. In J. W. van Prooijen & P. van Lange (Eds.), *Cheating, corruption, and concealment: The roots of dishonest behavior.* Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9781316225608.017.

*Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–60). New York: Academic. https://doi.org/10.1016/s0065-2601(08)60369-x.

CHAPTER 17

Formulaic Sequences as a Potential Marker of Deception: A Preliminary Investigation

Samuel Larner

Previous research into deception detection has argued that deception is more cognitively demanding than truth-telling. This additional cognitive load can lead to changes in linguistic and non-linguistic behavior, which in turn can be considered cues to deception (e.g., DePaulo et al., 2003). While the majority of deception research is rooted in psychology, this research approaches deception from a linguistics perspective by proposing and empirically testing a feature of language used to manage the cognitive demands of interpersonal communication. Since formulaic sequences—an umbrella term for sequences of words including metaphors, clichés, collocations, and routinized phrases—are stored holistically as single lexical items, they make the act of producing language less cognitively demanding. It is therefore hypothesized that individuals may seek to compensate for the additional cognitive demands of lying by increasing their reliance on formulaic sequences. To test this assertion, formulaic sequences were identified in a corpus of 1600 deceptive and truthful hotel reviews (Ott, Choi, Cardie, & Hancock, 2011), totaling 239,113 words, using an automated procedure based on a specially compiled dictionary of formulaic sequences. The results shed light on the relationship between formulaic sequences and deceptive language, their potential role in detecting deception, and the generalizability of findings to other types of texts.

S. Larner (✉)
Department of Languages, Information and Communications,
Manchester Metropolitan University, Manchester, UK
e-mail: S.Larner@mmu.ac.uk

Introduction: Formulaic Sequences Reduce Cognitive Load

Research has established that the act of telling lies is more cognitively demanding than telling the truth (e.g., Buller & Burgoon, 1996; DePaulo, 1992; DePaulo et al., 2003; Zuckerman, DePaulo, & Rosenthal, 1981). Vrij, Granhag, Mann, and Leal (2011) provide six reasons why this may be the case:

1. A lie firstly needs to be formulated, which then needs to be monitored to ensure it is plausible and in keeping with what the observer knows or may come to know, and then the liar has to remember what they have said and to whom;
2. Liars are less likely to take their credibility for granted and so will try to convince the observer of their honesty more so than truth-tellers;
3. As a result of this, liars carefully monitor the observer(s) to try and establish if they are actually being convincing;
4. Liars have to consistently remind themselves to role-play;
5. Liars generally must not only maintain a lie, but also suppress the truth; and finally
6. Telling the truth may be automatic, but telling a lie is intentional and deliberate.

Previous investigations into verbal and nonverbal cues to deception have identified features which occur because of this excess cognitive load; in other words, verbal and nonverbal behaviors "leak" out as a consequence of increased cognitive load. Verbal cues to deception in particular have been the focus of much research, which is perhaps unsurprising given that language is the primary mechanism through which deceptive communication is transmitted and received (Galasiński, 2000). However, research to date into verbal cues to deception has predominantly been conducted from a social psychology perspective (with some notable exceptions, e.g., Galasiński, 2000; Shuy, 1998). This approach typically emphasizes the importance of language features but provides psychological explanations for why those features occur, related to the psychological state of mind of the deceiver. This chapter approaches deception detection from a different perspective—that of linguistics—and instead of focusing on verbal features related to leakage, the focus is instead on a linguistic mechanism underlying communication more generally: the phenomenon of formulaic sequences. By exploring the underlying mechanisms of language production, it may be possible to identify a linguistic strategy for coping with deception, which may be more universally applicable, as opposed to previous investigations which have, in effect, identified side effects of deception. After firstly exploring the nature of formulaic sequences, this chapter will report the results of an empirical investigation to determine the utility of formulaic sequences for deception detection.

The opportunity for novelty in language is vast:

> There is no doubt that essentially all speakers of a language are free to produce sentences they have never heard or produced before. Very few people, on seeing two blue rabbits in a fish-bowl, are going to be poorly equipped, linguistically, to express their experience, even though the sentence they would need to create for the task would undoubtedly be completely novel to them. (Fillmore, 1979, p. 95)

However, whilst the potential to produce novel utterances is limitless, speakers appear "to renounce the great freedom that the language offers" (Corrigan, Moravcsik, Ouali, & Wheatley, 2009, p. xiii). Nattinger and DeCarrico (1992) suggest that "just as we are creatures of habit in other aspects of our behaviour, so apparently are we in the ways we come to use language" (p. 1) and we have "preferred formulations" for expressing ideas (Wray, 2006, p. 591). This routinization of language can accurately be characterized as *formulaic,* which Wray (2002) defines as "[w]ords and word strings which appear to be processed without recourse to their lowest level of composition" (p. 4). In other words, just as the mental lexicon contains single lexical items, so too does it contain sequences or chunks of words, which are also processed holistically as single lexical items. Many terms exist for describing different aspects of formulaicity, including idioms, clichés, and metaphors as well as more specialized, technical definitions, such as phrasal lexemes (e.g., Moon, 1998), recurrent phrases (Stubbs & Barth, 2003), and situation-bound utterances (e.g., Kecskés, 2000), to name just a few. In fact, Wray (2002) found 57 different terms each describing something that can be thought of as formulaic and therefore proposed the term *formulaic sequence* as a way to unify a disparate field of research:

> a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (p. 9)

The key point, for present purposes, is the idea that sequences of words are retrieved whole rather than the words that make up a formulaic sequence being analyzed individually. This is what Wray (2002) calls *needs-only analysis*; that we only break down and analyze sequences of words if some need arises. In this way, needs-only analysis accounts for irregularity in formulaic language. Phrases and sequences of words which, if analyzed, would be found to contain obsolete vocabulary and ungrammatical structures, do not cause problems in daily interaction precisely because "they do not invite analysis" even though they could be analyzed if analytical processing were activated (p. 131). Wray (2002) provides the example of the formulaic phrase *by and large* to illustrate her point:

> The word *large* in *by and large* is not associated with the regular word meaning 'big' because there is no demand on native speakers ever to analyze the phrase and assign a meaning to its component parts. Its meaning and functions are stand-alone, so no analysis is necessary. (p. 132, original emphasis)

Clearly, there is a processing advantage to using prefabricated sequences of words wherever possible: it is cognitively more economical to use formulaic sequences than it is to produce entirely novel language. It stands to reason that if formulaic sequences offer a processing advantage in everyday language use, their significance should become more apparent in contexts where cognitive load is further increased. Being deceitful presents one such context. As such, considering that being deceptive is cognitively demanding, and given that formulaic sequences offer a processing advantage, this chapter aims to investigate whether reliance on formulaic sequences increases during deception compared to telling the truth.

It is impossible to say exactly how much language is produced formulaically although there have been some attempts at quantification. Erman and Warren (2000) claim that 55% of spoken and written language may consist of *prefabs*, which they define as "a combination of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization" (p. 31). They provide examples of prefabs such as *out of date*, *at the time*, *in the end*, *here and there*, *a waste of time*, *for some reason* and *all over the place*. Pawley and Syder (1983) argue that "the largest part of the English speaker's lexicon consists of complex lexical items including several hundred thousand lexicalized sentence stems" (p. 215) which they define as "a unit of clause length or longer whose grammatical form and lexical content is wholly or largely fixed" (p. 191). Examples provided by Pawley and Syder (1983) include *it's on the tip of my tongue*, *some people are hard to please*, *call me after work*, *would you like some more?* and *speak for yourself*, again to provide only a representative few. The reason for a lack of certainty regarding the exact proportion of formulaic language compared to novel language in everyday language use is a consequence of differences in how formulaic language is measured, how it is identified, and the types of data that have been evaluated (with some being highly conventionalized and context sensitive). Nonetheless, it is widely accepted that formulaic language is ubiquitous and prevalent in language (Wray, 2002). As such, if it can be demonstrated that formulaic sequence usage does increase with deception, then this potential new marker of deception should be relatively easy to find in deceptive texts, further making it an ideal candidate for investigation. The following section outlines the data used in this research, before then considering the most appropriate way to identify formulaic sequences for this exploratory investigation.

DATA: THE DECEPTIVE OPINION SPAM CORPUS

To investigate the potential increase of formulaic sequence usage in deceptive texts, it was necessary to analyze a corpus of truthful and deceptive language. The Deceptive Opinion Spam Corpus (Ott, Cardie, & Hancock, 2013; Ott et al., 2011), a publicly available corpus, was identified as the most appropriate. This is on the basis that the corpus is large, meaning that the results are potentially representative of truthful and deceptive written language. Secondly, the texts are written by individuals in isolation. This means that there is no influence from other writers or speakers that may have primed formulaic sequences. Finally, as a general corpus of truthful and deceptive texts, the participants did not know that formulaicity was under investigation and so their use of formulaic sequences cannot have been primed by the research context.

The Deceptive Opinion Spam Corpus consists of 800 truthful and 800 deceptive written reviews of the 20 most popular hotels in Chicago, US. The truthful reviews were collected from six online travel websites (Expedia, Hotels.com, Orbitz, Priceline, TripAdvisor, and Yelp). The criteria for inclusion were that the reviews had to be written in English and had to be longer than 150 characters. Furthermore, the reviews had to be written by authors who had produced more than one review, on the basis that these may be more likely to be genuine. Ott et al. (2011, 2013) further subdivide the corpus into positive and negative affect, with five-star reviews indicating a truthful positive experience and one- and two-star reviews indicating a truthful negative experience of the hotel (400 reviews in each condition). For the present purposes, affect is not being considered and so the corpus is being treated as a whole.

Authors were recruited to write the deceptive hotel reviews through Amazon Mechanical Turk. Amazon Mechanical Turk is a crowdsourcing service. This enables small tasks (such as writing hotel reviews as in this instance) to be completed relatively cheaply by "anonymous online workers (known as *Turkers*)" (Ott et al., 2011, p. 311). To this end, they recruited a pool of 800 Turkers all based in the US, who were required to write a deceptive hotel review for one of the 20 hotels (Turkers were allocated evenly across the 20 hotels). They were given 30 minutes to complete the task and were paid one US dollar for their participation. Specifically, Turkers were presented with the name and website for a specific Chicago hotel and were told to assume that they worked for that hotel's marketing department. They were further instructed to pretend that their employer required them to write a fake review from the perspective of a customer, which would be posted on a travel review website. As such, Turkers were informed that the reviews needed to sound realistic about the hotel (Ott et al., 2011). As with the truthful reviews, participants were instructed to write about the hotel either positively or negatively, but again, affect is not being considered here.

The total corpus consists of 239,030 words, in which the shortest text is just 25 words whilst the longest is 784. The average length of texts is 149 words. Whilst, from a research perspective, the discrepancy in text lengths may be problematic, from a linguistics point of view, introducing text length as an uncontrolled variable introduces some resilience, and indeed, much needed ecological validity to the results.

METHOD: IDENTIFYING FORMULAIC SEQUENCES

Whilst there has been a great deal of empirical research into the phenomenon of formulaic sequences, the biggest challenge comes from identifying genuine examples in text, and an inherent difficulty with researching formulaic sequences is that there is a lack of consensus over how best to identify them. Indeed, Wray (2008) comments that "[i]dentifying formulaic sequences in normal language can be rather like trying to find black cats in a dark room: you know they're there but you just can't pick them out from everything else" (p. 101). Similarly, Erman and Warren (2000) caution that identifying every instance of a formulaic sequence in a text, and only genuine instances of formulaic sequences, "is in practice impossible" (p. 33). This stems from the fact that what is formulaic for some people may not be formulaic for others. Furthermore, formulaic sequences can vary in the extent to which they are fixed and they do not always have firm borders, with Wray (2008) highlighting that "the problem with formulaic language is that between the extremes of what is *definitely* formulaic and what is definitely *not* formulaic, there is a sizeable amount of material that may or may not be" (p. 93, original emphasis). Different techniques for identification can cope with various types of formulaic sequences, but no single technique which can identify them all has yet been developed. On this point, then, the main ways through which formulaic sequences can be identified include:

- the researcher's own intuitions (what "feels" formulaic to the researcher);
- consensus derived from speech communities (whether a group of people from the same speech community can agree on a sequence being formulaic);
- structural analysis (whether a literal interpretation is or is not possible, the degree to which word order can be changed, and whether lexical insertions, inflections, and replacements are possible);
- pragmatic and functional analyses (whether the candidate formulaic sequences are tied to particular acts, events, and/or rituals or whether they are tied to specific functions);
- established reference lists and dictionaries; and
- statistical analyses (in which more frequently occurring sequences are considered formulaic, and rarer sequences are considered less formulaic).

These different techniques represent the distinction between a rigorous but narrow approach (such as statistical analyses) and broader, but less reliable approaches (such as using intuition) and there are clear advantages and problems with each of these techniques. Whilst reliability may be low, intuition can be used for identifying formulaic sequences on the premise that native speakers recognize formulaic language as having special status (Van Lancker-Sidtis & Rallon, 2004), and on the basis that as members of their own speech communities, researchers "often are the self-appointed arbiters of what is idiomatic or formulaic in their data" (Wray, 2002, p. 20). It is possible to improve reliability by using second-raters or panels of independent judges and "there should be a certain resilience in a consensus achieved in this way" (Wray, 2002, p. 22). By contrast, identifying formulaic sequences by reference to a dictionary or reference list offers a more reliable approach than intuition, but being reliant on automated processes, requires exact matching between the reference list and formulaic sequences in the data. In other words, whereas intuition can be useful for identifying formulaic sequences with unclear borders or variable lexical items, automated approaches are faster, more reliable, and require just one researcher (rather than a panel of judges), but are limited to identifying only exact matches. Wray (2008) cautions that despite reference lists often being used as alternatives to intuitions, researchers still need to understand how the lists were compiled, and warns against using lists which have "gained authority simply by virtue of being published" (p. 109).

Larner (2014) highlights the importance of three issues when considering the identification of formulaic sequences specifically in forensic contexts: validity (what was identified is the same as what was intended), reliability (analyses repeatedly produce the same results), and feasibility (how well the method can be applied to forensic data, taking into account the often-limited time and resources). Larner (2014) set out to establish whether authors' texts can be differentiated for forensic authorship analysis purposes, based on the proportion of formulaic sequences that each author uses and to investigate this, he proposed a method for identifying formulaic sequences that satisfied—as far as possible—these three criteria. He constructed a reference list of formulaic sequences informed by a multitude of different web sources to ensure representativeness of a large speech community. By using the web, he argued that it is possible to harness the power of speech community intuitions, but in such a way that enables reliability.

The reference list used in the current research is the same as that developed by Larner (2014). Through searching the web for terms generally accepted as formulaic (e.g., *list of proverbs*, *list of clichés*, *list of common phrases*, *list of similes*, *list of popular sayings*), it was possible to collect 17,973 formulaic sequences. For each search term, only the first five pages of results were explored, meaning that these sites were those deemed to be the most popular. The websites ranged in their purposes: the majority aimed to help speakers of

other languages sound more native-like in English, whilst others provided the etymology of these entries. The list of entries was then processed to ensure consistency across entries and to improve reliability:

1. All pronouns were replaced with an asterisk wild card to allow for pronoun variation in formulaic sequences, e.g., ** bark is bigger than * bite* enables a range of variants such as *his bark is bigger than his bite* and *her bark is bigger than her bite* to also be identified.

2. Wray ([2002](#)) explains that sometimes simply starting an idiom can be sufficient for it to be recognized, rendering the need for writing the complete phrase obsolete. Since longer stretches of text are less likely to be matched in their exact form, the decision was made to also include shorter variants, e.g., *a rose by any other name would smell as sweet* being shortened to *a rose by any other name*, which was then added alongside the longer entry on the list, offering the opportunity to match either the longer or shorter version.

3. Given that identical character matching is so crucial to the identification process, punctuation was potentially problematic since it can be variable, depending on the level of formality of the text (i.e., whether the author adheres to standard punctuation or not). As such, variants were included in the list with punctuation marks removed.

4. Given the nature of the web, it is not possible to determine in which country the websites were hosted, nor the native language of their authors. However, there were many instances of North American spelling variants. This is not necessarily problematic since the Deceptive Opinion Spam Corpus contained texts believed to be written by North American, native English speakers. However, when originally developing this reference list, Larner's ([2014](#)) participants were native British English speakers, and so British English variants were added, e.g., *good fences make good **neighbours**, in **honour** of*, and *in self-**defence***.

5. There were many duplicates in the list, which adds further support to the argument that these were considered to be formulaic for the speech community at large rather than any one individual. Duplicates were removed.

The final reference list contained 13,412 formulaic sequences. Of course, whilst the list exists in one file, and the data exists as another set of files, software is needed which can read the reference list and automatically identify matches in all of the data. AntConc v.3.4.4 (Anthony, [2016](#)) was identified as the most appropriate for this purpose. After importing the data to AntConc it is possible, using the "advanced" features, to load a file containing search terms (in this case, the individual formulaic sequences), which are then identified in the data. Table [17.1](#) provides a range of examples (in bold) that illustrate the advantages and disadvantages of this identification process.

Table 17.1 Examples of formulaic sequences identified using the automated reference list approach

Extract	Examples of formulaic sequences identified in the data
1	Quiet tree-lined streets, real **old-school** townhouses (instead of ugly towering monstrosities that are twenty-first century condominiums)
2	Then again Delaware St is not exactly the widest street **in town** and there are buildings **in front of** us, so can't have it all! Yet one thing that really I do get a service that is **a little bit** more personalized
3	The lady **on the phone** was anything but courteous and competent
4	I had to explain myself **all over** again (my **pet peeve** when it comes to **on-the-phone** customer service)
5	**Case in point**—he had to send the slip of paper on which I agree to the nightly rate
6	I really wasn't expecting hotels as small as Talbott to have a 24-hour room service, but **lo and behold**, they do
7	Cute boutiques (instead of Gap, **Banana Republic** and other mainstream mass-produced brands)
8	Enough space to walk around, do **a chicken** dance, and strut my Herve and Louboutins in

It can be seen from Table 17.1 that in extracts 1, 2, 3, 4, 5, and 6, a range of formulaic sequences have been correctly identified (e.g., *old-school*, *in town*, *in front of*, *a little bit*, *on the phone*). It is also noteworthy that a variant form of *on-the-phone* in extract 4 (i.e., *on the phone*) has been identified in extract 3, which demonstrates the value of editing the reference list to include variants of some formulaic sequences that contain different punctuation marks, since this instance could have been missed. Extract 7 (*Banana Republic*) illustrates a limitation of this method: whilst *Banana Republic* does exist in the formulaic sequences reference list (which can be glossed as "a politically instable country whose economy relies on exporting just one limited resource"), in this specific example, it is being used in a different sense, with *Banana Republic* referring to the name of a clothing brand. However, it is possible that *Banana Republic*, used in this sense, is still formulaic since speakers may not necessarily analyze the individual words and instead treat the name of the clothing brand holistically (i.e., formulaically). In extract 8, *a chicken* has been identified. This formulaic sequence exists in the reference list (with the meaning "cowardly"), where here it is used to refer to a particular type of dance. Arguably, *a chicken dance* does have potential to be formulaic although the key point is that this method sometimes identifies formulaic sequences which are not being used formulaically in a particular context. Finally, there are some instances that "feel" formulaic to me, as the researcher, but which have not been identified. For instance, in extract 2, I would have liked *then again*, *not exactly the*, *you can't have it all*, and *one thing* to be identified, whilst in extract 6, *room service* seems important. Likewise, the reader may too identify other sequences of words that intuitively seem formulaic. However, as identified earlier, the intuitions

of one person are less reliable, so instances like these have to be accepted as legitimately not identified in the data.

<div align="center">RESULTS</div>

All 1600 texts in the Opinion Spam Corpus were used in the analysis and a total of 2380 formulaic sequence tokens were identified. It became apparent that some instances of matches were being used in a sense not originally captured by the reference list (cf. extract 7 in Table 17.1). For instance, the sequence *Take the Cake* was identified, but referred to a specific bakery rather than the idiomatic meaning of the sequence ("outstanding in some respect, either for being extremely positive or extremely negative"). Likewise, the formulaic sequence *four seasons*, typically glossed as "somebody who does something all year round," was used exclusively as a proper noun for the hotel chain in these data. The phrase *at the Best* in the context of *at the Best Western* [hotel] was also highlighted as a match for the formulaic sequence *at * best*. In the formulaic sequence *at * best*, the asterisk wildcard was intended to allow pronoun variation to capture variations such as *at my best*, *at your best*, *at her best* (meaning "of the highest standard possible"). However, in these instances, rather than identifying formulaic variants of *at * best*, the name of the hotel chain *at the Best* was identified. Furthermore, the majority of these contentious matches were so linked to the hotel industry (i.e., names of hotels) that if they remained in the analysis, the results would not be replicable on other types of data. As such, the decision was made to remove all the clearly non-formulaic matches (debatable entries such as *a chicken*, discussed above, remain in the analysis). A total of 101 formulaic sequence tokens were removed, leaving 2279 formulaic sequences being identified across all the data. The 2279 formulaic sequence tokens were made up of 525 different formulaic sequence types. The identified formulaic sequences ranged in length from one to six words as exemplified in Table 17.2. In total, the 2279

Table 17.2 Examples of identified formulaic sequences ranging in length from one to six words

Number of words constituting formulaic sequence	*Matches*	*Examples*
1	276	okay, on-the-phone, plain-as-day, state-of-the-art
2	1260	above average, final straw, in future, no brainer
3	504	blew me away, down to earth, in my opinion, in the meantime
4	217	bump in the road, icing on the cake, spur of the moment, set my sights on
5	16	a piece of my mind, as hard as a rock, at our beck and call, bad taste in my mouth
6	6	you get what you pay for, to make a long story short, cost an arm and a leg

formulaic sequence matches constituted 5292 individual words, meaning that each formulaic sequence was, on average, 2.3 words in length.

By way of comparison, Larner (2014) also found that through using this reference list with short personal narratives as data, the two- and three-word formulaic sequences were again the most frequently identified, with formulaic sequences having a mean length of 2.6 words. This suggests either that the most commonly used formulaic sequences tend to be shorter, or that this particular method for identification is more suited to the shorter forms of formulaic sequences. The latter seems more plausible in light of the fact that as part of the reference list editing, clipped forms of longer formulaic sequences were included to increase the likelihood of identification.

The next stage is to determine whether there is actually a difference in the use of formulaic sequences between the truthful hotel reviews and the deceptive hotel reviews, as originally hypothesized. However, an important methodological decision must firstly be made: whether to count the number of matches identified in the data, or whether to count the number of words that make up a formulaic sequence. From a theoretical perspective, it is sensible to count the number of matches (i.e., the total number of formulaic sequences identified). Since formulaic sequences are stored as holistic chunks of words, all the words that comprise a single formulaic sequence are, in principle, stored and analyzed as a single word and should therefore be counted as a single word. However, this is problematic in practice due to the fact that very short texts—the shortest of which is 25 words—are being considered. The number of matches in these texts may be so low that they restrict any meaningful analysis. As such, in line with Larner (2014), the number of words that make up formulaic sequences was taken as the measure. For instance, *bump in the road* is counted as four formulaic sequence words rather than one formulaic sequence hit. Whilst it is true that there can be great variation between the length of formulaic sequences, Table 17.2 demonstrates that there were relatively few of the five- and six-word formulaic sequences and nothing longer. Likewise, the mean average length of identified formulaic sequences was 2.3, and the median and mode were both 2. Furthermore, the standard deviation was low ($\sigma = 0.851$) indicating that for present purposes, it is acceptable to use the count of words that make up formulaic sequences as the measure. This means that a raw score of formulaic language words can be calculated against the words that have not been identified as formulaic in each text. A further consideration is that the lengths of texts in the corpus did vary, and so a normalized count of formulaic words per 100 words was calculated rather than a raw frequency count. By calculating the normalized count, it will be possible to make claims about whether the language used in deceptive reviews is more or less formulaic than that of truthful reviews.

Since the data significantly deviates from a normal distribution (Genuine: $W = 0.857$, $p < 0.01$, Deceptive: $W = 0.839$, $p < 0.01$), the non-parametric Kruskal–Wallis test was used. A Kruskal–Wallis H test showed that there was not a statistically significant difference in formulaic sequence words between

Table 17.3 Ten most frequent formulaic sequences identified in the truthful and deception sub-corpora

Truthful	Deceptive
check in (67)	my husband and I (80)
check-in (63)	check in (76)
housekeeping (50)	check-in (36)
check out (32)	as if (35)
go back (32)	go back (33)
at night (28)	at least (30)
at least (25)	housekeeping (29)
within walking distance (25)	at night (27)
on business (19)	check out (23)
my husband and I (18)	at home (22)

the truthful and deceptive hotel reviews, $X^2 = 3.020$, df $= 1$, $p = 0.82$, with a mean rank formulaic words score of 780.80 for genuine hotel reviews and 820.20 for deceptive hotel reviews. We can therefore interpret this as meaning that whilst formulaic language words do occur more frequently in deceptive hotel reviews, the difference in means between the deceptive and truthful hotel reviews is not significant. Table 17.3 shows the top 10 formulaic sequence types that were identified in the truthful and deceptive data (numbers within parentheses indicate the frequency with which they occur).

It is interesting to notice that in the truthful data, the formulaic sequence *check in* is the most frequent, whilst in the deceptive data it is *my husband and I*. By contrast, *my husband and I* is only the 10th most frequently used formulaic sequence in the truthful data. Likewise, *housekeeping* is used more frequently in the truthful data compared to the deceptive data, and the formulaic sequences *within walking distance* and *on business* occur in the top 10 for truthful data, but not for deceptive data (incidentally, these two formulaic sequences were ranked 13th and 12th, respectively in the deceptive data). It can also be seen that the deceptive data contains two formulaic sequences that do not occur in the top 10 most frequently used formulaic sequences in the truthful data: *as if* and *at home*. From these top 10 formulaic sequences, then, a total of 12 types are used, as demonstrated in Table 17.4.

A chi-square test of independence was performed to examine the relation between the remaining eight formulaic sequences and deception. The relation between these variables was significant, X^2 (11, $N = 750$) $= 196.673$, $p < 0.01$. However, since four of these formulaic sequences had a frequency of zero in either the truthful or deceptive data, they could have potentially skewed the results. Therefore, to further improve resilience of these results, these four formulaic sequences were removed. A chi-square test of independence was performed to examine the relation between the remaining eight formulaic sequences and deception. The relation between these variables was significant, X^2 (7, $N = 649$) $= 3839.011$, $p < 0.01$. This indicates that despite the previous finding that the overall count of formulaic words

Table 17.4 Frequency of occurrence across the truthful and deceptive sub-corpora for the twelve most frequent formulaic sequence types

Formulaic sequence	Frequency in truthful	Frequency in deceptive
check in	67	76
check-in	63	36
housekeeping	50	29
check out	32	23
go back	32	33
at night	28	27
at least	25	30
within walking distance	25	0
on business	19	0
my husband and I	18	80
as if	0	35
at home	0	22

to non-formulaic words does not differ between truthful and deceptive hotel reviews, the frequency with which individual types of formulaic sequence occur in both truthful and deceptive data does differ.

Frequent Clusters as Candidate Formulaic Sequences

So far, we have seen that whilst deceptive hotel reviews do contain more formulaic words than truthful reviews, this difference is not statistically significant. It has also become evident that some formulaic sequences are used more frequently in the deceptive hotel reviews than the truthful hotel reviews. But the analysis so far assumes, as was argued by Larner (2014), that the reference list was in fact valid as a way to identify formulaic sequences. Both Wray (2002) and Read and Nation (2004) independently assert that none of the various techniques for identifying formulaic language alone are adequate, and that valid results can only be obtained through using more than one form of analysis. To this end, triangulation is likely to produce findings which are more valid and reliable. The final stage of the analysis therefore adopts a statistical approach in order to identify sequences of words that may be considered to be formulaic by virtue of the frequency with which they occur. Using *WordSmith Tools* 7.0 (Scott, 2017), the 20 most frequently occurring clusters were extracted from the truthful and deceptive sub-corpora, as indicated in Table 17.5.

When comparing the list of the top 20 clusters identified in the truthful and deceptive sub-corpora, it is evident that the majority of clusters in these lists are shared between both sets of reviews. However, some clusters do occur only in the truthful sub-corpora (*on the*, *room was*, *for the*, *from the*, *was very*), whilst some clusters occurred only in the deceptive reviews (*I had*, *in Chicago*, *when I*, *to be*, *I would*). Despite these obvious differences, the main

Table 17.5 20 most frequently occurring clusters in the truthful and deceptive sub-corpora

Rank	Truthful clusters	Frequency	Deceptive clusters	Frequency
1	in the	651	at the	647
2	the hotel	593	in the	603
3	of the	529	of the	544
4	at the	476	the hotel	535
5	the room	447	I was	532
6	and the	438	and the	485
7	to the	392	the room	450
8	on the	330	and I	403
9	it was	314	it was	402
10	this hotel	305	this hotel	383
11	room was	294	to the	346
12	for a	281	I had	301
13	we were	275	in Chicago	265
14	I was	270	when I	253
15	for the	248	for a	249
16	from the	233	to be	243
17	was very	230	the staff	242
18	the staff	208	I would	238
19	and I	207	was a	238
20	was a	204	we were	238

difference is the frequency with which these clusters occur. For instance, the most frequent cluster in the truthful reviews (*in the*) occurs 651 times, and 603 in the deceptive sample, making it only the second most frequently occurring cluster. Likewise, the most frequent cluster in the deceptive reviews (*at that*) occurs 647 times, but only 476 times in the truthful sub-corpus, making it only the fourth most frequent cluster. Therefore, it is necessary to determine whether these frequencies of occurrence differ significantly between the truthful and deceptive reviews. Table 17.6 provides all of the clusters that were identified as the 20 most frequent clusters between the two sub-corpora, resulting in a total of 25 clusters.

A chi-square test of independence was performed to examine the relation between the top 25 frequently occurring clusters and deception. The relation between these variables was significant, X^2 (24, $N = 14,522$) = 3839.011, $p < 0.01$, indicating that the frequency with which these individual clusters occur is significantly different between truthful and deceptive hotel reviews. However, it can be seen in Table 17.5 that ten of those clusters occurred within the top 20 most frequent clusters in one sub-corpus but not the other, leading to a frequency count of zero in one of the sub-corpora. Therefore, to further test the data, these ten clusters (*on the*, *room was*, *for the*, *from the*, *was very*, *I had*, *in Chicago*, *when I*, *to be*, *I would*) were excluded from the analysis. A chi-square test of independence was performed to examine the relation between the top 15 frequently occurring clusters and deception. The relation between these variables was also significant, X^2 (14, $N = 11,887$) = 175.159, $p < 0.01$. This result indicates that the frequencies with which these

Table 17.6 The 20 most frequent cluster types in the truthful sub-corpus and deceptive sub-corpus represent 25 cluster types overall

Cluster	Frequency in truthful hotel reviews	Frequency in deceptive hotel reviews
in the	651	603
the hotel	593	535
of the	529	544
at the	476	647
the room	447	450
and the	438	485
to the	392	346
on the	330	0
it was	314	402
this hotel	305	383
room was	294	0
for a	281	249
we were	275	238
I was	270	532
for the	248	0
from the	233	0
was very	230	0
the staff	208	242
and I	207	403
was a	204	238
I had	0	301
in Chicago	0	265
when I	0	253
to be	0	243
I would	0	238

specific clusters occur, whether they are actually formulaic or not, are related to whether the hotel review was deceptive or truthful.

DISCUSSION

The results presented above provide three key findings about the relationship between formulaic sequences and deception: (1) the reliance of formulaic sequences does not appear to increase with deception as originally predicted; (2) the frequency with which specific types of formulaic sequences are used does differ between truthful and deceptive hotel reviews; and (3) analysis of frequent clusters, as a different way to identify formulaicity, confirms the finding of (2). Each of these findings will now be discussed.

The first area that warrants discussion is the fact that the proportion of formulaic sequences did not increase in the deceptive hotel reviews compared to the truthful reviews. There are three possible explanations for this. The first is that the method for identifying formulaic sequences was not sufficiently robust. However, it should be borne in mind that Larner (2014) applied the same method to different data and was able to identify clear differences between authorial writing styles. The second possible explanation is

that this finding is accurate—formulaic sequences are simply not a response to the cognitive demands of lying. Whilst this is of course a legitimate possibility, such a finding would fly in the face of the research which demonstrates formulaic sequences to be a linguistic strategy for coping with the cognitive demands of producing language. To take such a stance based on the results of a relatively small, exploratory piece of research would be unsound. This leaves the third possible explanation, that formulaic sequences may well still be a cue to deception, but the data used in this investigation was deficient in some areas, meaning it did not draw out these differences significantly enough. The premise of this paper is that deception is cognitively demanding and that producing language is also cognitively demanding, for which formulaic sequences offer a solution. As such, the assumption has been made that (a) the very act of writing the hotel reviews was cognitively demanding and that (b) the effort to produce convincing deceptive hotel reviews was sufficiently cognitively taxing. In the description of the Deceptive Opinion Spam Corpus, Ott et al. (2011) indicate that this may not necessarily be the case.

The main area of contention for the present purposes is Ott et al.'s (2011) finding that 12% of the participants in the Deceptive Opinion Spam Corpus (i.e., those producing deceptive hotel reviews) uploaded their review in less than a minute, despite having 30 minutes to complete the task. They further found that there was no correlation between the time taken to complete the review and the mean length of the review (indeed, the quickest review was completed in just five seconds but contained 114 words). Ott et al. (2011) argue that a likely explanation is that the Turkers firstly reviewed the task to determine what would be involved, then wrote the review, and then accepted the task and immediately uploaded their pre-written response. This does not affect the quality of the deceptive review, but for present purposes, it is hard to argue that the data was written under cognitively demanding circumstances, if some participants wrote at their leisure. This might therefore explain why there is a qualitative difference in the formulaic sequences identified, but not the overall proportion of formulaic sequences identified—the cognitive demand simply was not high enough to activate formulaic sequences as a coping mechanism.

Furthermore, it should be borne in mind that the task of writing a deceptive hotel review under experimental conditions has to be considered a very low-stakes task. The Turkers were offered one dollar for their participation and potentially had little investment in whether their review was convincing or not. The decision to use this data was motivated by the fact that an experimental, heavily constrained approach is the most suitable for exploratory research. The findings should therefore be considered justification for replicating the study on other, more cognitively demanding forms of data. In this regard, real-world authentic data, such as police witness statements or interviews would be ideal, since these can be considered to be genuinely high-stakes situations.

Moving on, it is important to consider why the types of formulaic sequences used would differ between truthful and deceptive hotel reviews. Why is it that the deceptive reviewers used the formulaic sequences *my husband and I* four times more frequently than truthful reviewers? Why would *within walking distance* and *on business* occur more frequently in truthful reviews than in deceptive reviews, whilst the reverse is true for *as if* and *at home*? There are two possible explanations for this. Potentially, members of a speech community are socialized to recognize the pragmatic and functional aspects of formulaic sequences. The focus in this research has been on the psycholinguistic aspect of formulaic sequences, but there is also an important sociolinguistic dimension. It is possible, through repeated exposure to forms of deception (e.g., through drama, film, and television) that the Turkers had come—at a purely intuitive, subconscious level—to recognize that some formulaic sequences are functionally appropriate when being deceptive. Although not attested in this data, an example might be the urge to utter *you can trust me*, or *are you calling me a liar*?, when being deceptive: formulaic sequences which liars may draw on to convince their interlocutor that they are believable. Whilst clear examples of this type have not been identified in the data, this presents an intriguing linguistic explanation for the difference. Alternatively, the formulaic sequences that were identified may belie something about the psychological state of mind of the deceiver. For instance, the deceptive reviewers' reliance on *my husband and I* may be a distancing strategy, whilst the truthful reviewers' frequent use of *on business* stands in direct contrast to the deceptive reviewers' frequent use of *at home*, as if playing on the real-world experiences of the reviewers, on which their reviews are based. Perhaps those producing the deceptive reviews anchored what they wrote in their personal, actual experiences of using hotels for pleasure rather than business, so their ability to construct an entirely novel lie, disconnected from reality, was limited. Of course, this goes entirely beyond the realms of what a linguist would conclude, so I offer this not as an explanation, but rather as a way to illustrate the way that linguists and psychologists might fruitfully work together in exploring deception from our different vantage points. The bigger finding, perhaps, is that these results indicate that with further testing on far more diverse types of data, it may be possible to extract a list of formulaic sequences that are more indicative of either truth or deception.

Finally, it is noteworthy that both the formulaic sequences reference list, and the frequent clusters approach, revealed differences in the types of formulaic sequences used by truthful and deceptive reviewers. Given the challenges discussed in the methods section (namely, that it is impossible to identify every instance), it is pleasing that by identifying formulaic sequences in two very different ways, the same result has been found: that formulaic sequence types differ between truthful and deceptive reviews. It is less pleasing, however, that formulaic sequences identified through the reference list approach were not simultaneously identified through the clusters approach.

On the one hand, this is less intuitively satisfying since successful triangulation would see the same formulaic sequence being identified through two or more approaches. On the other hand, regardless of the fact that different sequences have been identified in each case, the actual types used do differ, adding further support to the argument presented above, that maybe something at the psychological or sociolinguistic level motivates differences in how truth-tellers and deceivers wrote their hotel reviews.

Conclusion

The aim of this chapter has been to explore the use of formulaic sequences in deceptive and truthful data, with a view to determining whether there is a potential relationship between the use of formulaic sequences and deception. The key finding has been that whilst the proportion of formulaicity (measured as the total number of words that constitute a formulaic sequence compared to the number of novel words in each text) is marginally higher in deceptive reviews, this difference is not statistically significant. However, the linguistic types of formulaic sequences that are used between the truthful and deceptive hotel reviews do occur with different frequencies, and these differences are significant, even when the biggest differences are eliminated from the analysis. As such, I would like to conclude this chapter with both a recommendation and a suggestion. The recommendation is that further testing into the relationship between formulaic sequences and deception is warranted, and that future studies would benefit from exploring different types of data. Those produced in high-stakes scenarios, and those where cognitive load is heightened (such as unplanned speech, rather than editable writing) are likely to be the most fruitful. Since language is the primary mechanism for deception, and since being deceptive relates to a state of mind that is different from a truth-teller, my suggestion is that linguists and psychologists might benefit from closer working relationships. To date, our respective endeavours have been largely independent, but our fields clearly have a lot to offer, and collaborating on deception detection research may provide exciting new areas for interdisciplinary research.

References

Anthony, L. (2016). *AntConc version 3.4.4.* Tokyo, Japan: Waseda University.

Buller, D., & Burgoon, J. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x.

Corrigan, R., Moravcsik, E., Ouali, H., & Wheatley, K. (2009). Introduction. Approaches to the study of formuale. In R. Corrigan, E. Moravcsik, H. Ouali, & K. Wheatley (Eds.), *Formulaic language: Distribution and historical change* (Vol. 1, pp. xi–xxiv). Amsterdam: John Benjamins Publishing Co. https://doi.org/10.1075/tsl.82.02cor.

DePaulo, B. (1992). Nonverbal behaviour and self presentation. *Psychological Bulletin, 111*(2), 203–243. https://doi.org/10.1037//0033-2909.111.2.203.

DePaulo, B., Lindsay, J., Malone, B., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118. https://doi.org/10.1037//0033-2909.129.1.74.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text, 20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29.

Fillmore, C. (1979). On fluency. In C. Fillmore, D. Kempler, & W. S.-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85–101). London: Academic Press. https://doi.org/10.1016/b978-0-12-255950-1.50012-3.

Galasiński, D. (2000). *The language of deception: A discourse analytical study*. London: Sage. https://doi.org/10.4135/9781452220345.

Kecskés, I. (2000). A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics, 32,* 605–625. https://doi.org/10.1016/s0378-2166(99)00063-6.

Larner, S. (2014). A preliminary investigation into the use of fixed formulaic sequences as a marker of authorship. *The International Journal of Speech, Language and the Law, 21*(1), 1–22. https://doi.org/10.1558/ijsll.v21i1.1.

Moon, R. (1998). Frequencies and forms of phrasal lexemes in English. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 79–100). Oxford: Oxford University Press.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Ott, M., Cardie, C., & Hancock, J. (2013). Negative deceptive opinion spam. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 497–501). Atlanta, USA: Association for Computational Linguistics.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. (2011). Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 309–319). Portland, USA: Association for Computational Linguistics.

Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191–226). New York: Longman.

Read, J., & Nation, P. (2004). Measurement of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 23–35). Amsterdam: John Benjamins Publishing Co. https://doi.org/10.1075/lllt.9.03rea.

Scott, M. (2017). *WordSmith tools version 7*. Stroud: Lexical Analysis Software.

Shuy, R. W. (1998). *The language of confession, interrogation, and deception*. London: Sage. https://doi.org/10.4135/9781452229133.

Stubbs, M., & Barth, I. (2003). Using recurrent phrases as text-type discriminators: A quantitative method and some findings. *Functions of Language, 10*(1), 61–104. https://doi.org/10.1075/fol.10.1.04stu.

Van Lancker-Sidtis, D., & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language & Communication, 24,* 207–240. https://doi.org/10.1016/j.langcom.2004.02.003.

Vrij, A., Anders Granhag, P., Mann, S., & Leal, S. (2011). Outsmarting the liars: Towards a cognitive lie detection approach. *Current Directions in Psychological Science, 20*(1), 28–32. https://doi.org/10.1177/0963721410391245.

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511519772.

Wray, A. (2006). Formulaic language. In E. K. Brown (Ed.), *The encyclopedia of language and linguistics* (pp. 590–597). Oxford: Elsevier. https://doi.org/10.1016/b0-08-044854-2/04777-5.

Wray, A. (2008). *Formulaic language: Pushing the boundaries.* Oxford: Oxford University Press.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology, 14,* 1–59. https://doi.org/10.1016/s0065-2601(08)60369-x.

The Reciprocal Nature of Lying and Memory: Memory Confabulation and Diagnostic Cues to Deception

Rachel E. Dianiska, Daniella K. Cash, Sean M. Lane and Christian A. Meissner

To the layperson, recalling from memory would seem fairly straightforward—if a person remembers an event, surely it must have occurred and will be reported accurately; however, the literature is replete with examples in which our own memory deceives us. For instance, could a person misremember having their aircraft shot down in Iraq? Such a distinctive event would be memorable to whomever experienced it. Yet, in 2015, a highly respected *NBC News* reporter, Brian Williams, retracted multiple, sometimes conflicting, narratives involving an aircraft that he was in (or an accompanying aircraft) coming under fire and being forced to make an emergency landing in Iraq. Later investigation revealed that the aircraft did not come under fire or incur damage, forcing Williams to recant his statement and receive a six-month suspension (Steel & Somaiya, 2015).

R. E. Dianiska (✉) · C. A. Meissner
Iowa State University, Ames, IA, USA
e-mail: dianiska@iastate.edu

C. A. Meissner
e-mail: cmeissner@iastate.edu

D. K. Cash · S. M. Lane
Louisiana State University, Baton Rouge, LA, USA
e-mail: daniellakcash@gmail.com

S. M. Lane
e-mail: slane@lsu.edu

The initial reaction to this recantation appeared to be disbelief that a news anchor would lie about something so easily verifiable. Rather than making a conscious decision to provide a wholesale fabrication, Williams' memory of the attack could have been created by smaller embellishments to the story that were enhanced through each retelling. Similar to fishing stories where the prized fish that was caught increases in size with each retelling, elaboration of the story over time may have allowed Williams to create a visually compelling memory that he would later recall and offer in future retellings, ultimately leading to a persuasive, albeit incorrect, memory of coming under fire.

Williams' story is one example of how the act of lying can influence a person's memory for an event or for the lie one tells. In this chapter, we first discuss research that documents how the act of lying can influence the content of liars' memories for the occasions when they lied and memories of the original experience. But an understanding of memory processes can be useful for uncovering deception as well. For instance, the content of memories of actual and fabricated events differ in characteristic ways, and people can be trained to utilize these features to discriminate between them. Furthermore, it is possible to magnify differences in the reports of liars and truth-tellers to increase detection. We will discuss these issues in the second half of the chapter.

The Influence of Lying on Memory

While the deception detection literature has primarily focused on the ability to detect false statements and accurately discriminate them from truthful statements (e.g., Bond & DePaulo, 2006), recent theoretical and empirical work has begun to examine whether the act of lying itself might influence one's memory for an event. Evidence suggests that the type of lie can both affect how well the act of lying is remembered, and whether a false memory is created (Dianiska, Lane, & Cash, 2016; Otgaar, Howe, Memon, & Wang, 2014; Otgaar, Howe, Smeets, & Wang, 2016; Polage, 2004; Polage, 2012; Vieira & Lane, 2013). Additionally, factors such as repetition of the lie (akin to Brian Williams' retelling of the Iraq narrative) can independently influence memory (e.g., Vieira & Lane, 2013). To avoid detection, a liar must remember that they lied as well as what they said while lying. Forgetting that one lied could lead people to later contradict themselves and thus get caught. In this section, we discuss two different types of lies (confabulations and denials), their potential effects on memory based on the source monitoring framework (SMF; Johnson, Hashtroudi, & Lindsay, 1993), and the available studies examining this interface between lying and memory.

Two Types of Lying—Confabulations and False Denials

Confabulations refer to lies that involve a person describing a specific event or experience as if it had occurred. As discussed above, it is plausible that

Brian Williams created small modifications to his story, perhaps based on his knowledge of other similar incidents that had occurred in Iraq. Of course, people create much more extensive fabrications, for instance, when a criminal describes his or her whereabouts for a period of time during which a crime occurred (i.e., an alibi). As we discuss in the next section, the fact that confabulations require constructive mental processes (although to varying degrees) has implications for how well these types of lies are remembered.

False denials refer to lies in which a person says that an event never occurred, despite the fact that the event truly did take place. This type of lie has the advantage of not requiring the liar to fabricate new details, and consequently, typically requires less effort to produce than a confabulation. Although there are a number of contextual variables that can potentially affect the choice of a lie type, the literature does suggest that certain forensic interviewing techniques may be more likely to elicit denials. For instance, accusatorial interviews in which suspects are confronted with accusations are likely to foster denials, both truthful and false, from interviewees (e.g., Vrij & Granhag, 2012). In contrast, inquisitorial interviews, that typically involve more open-ended questions, are more likely to elicit extended answers (see Meissner, Kelly, & Woestehoff, 2015 for an extended discussion of the differences between these interview types).

Confabulations and false denials differ to the extent that effortful, constructive mental processing is required. This difference affects the ability of liars to correctly ascertain whether they lied about a specific fact or event—source monitoring.

A Source Monitoring Perspective

The ability to determine the origin of a memory is called *source monitoring*. Examples of source monitoring include remembering whether your Uncle Bill or your father told you a particular story, whether you or a friend came up with an idea for a new business, or whether your memory of an event derives from actually being there or having imagined it. The source monitoring framework (SMF; Johnson et al., 1993) describes how such decisions are made. In this view, memories are composed of features that reflect the conditions under which they were encoded. On average, memories from different sources differ in the amount or type of features they contain, and this can be used to successfully discriminate them. For example, memories of perceived events (i.e., externally derived) tend to contain more perceptual (e.g., color, shape) and contextual (e.g., location and time) features, whereas memories of events that were imagined tend to contain features created by the act of constructing the image (e.g., cognitive operations), and fewer perceptual and contextual features. However, source monitoring decisions are not foolproof. Errors in source monitoring may arise when encoding is impaired in a way that limits the mnemonic cues (i.e., features) available at retrieval, or to the

extent there is similarity between two or more sources (see Lindsay, 2008). In general, the accuracy of source monitoring decisions is dependent upon the information available at the time of the memory decision, and the circumstances of the source monitoring judgment.

The source monitoring framework provides insight into how producing a confabulation or a false denial might influence one's later memory. Falsely describing an event, particularly if relatively elaborated, involves effortful constructive processing, and thus increases the availability of information about cognitive operations in the resulting memory trace. Note too, that an elaborated confabulation could also include described perceptual and contextual detail that would have been present if the liar had actually experienced an event. This latter information, if remembered at a later time, could mislead the person into thinking they had experienced the event (i.e., create a false memory). However, if the person also had access to information about cognitive operations, this would allow them to reject the notion that the event occurred because they remember creating the falsehood. Thus, the constructive process of generating a false description may ultimately prevent a liar from believing his or her own lies (Polage, 2004; Vieira & Lane, 2013). Of course, if the liar does not remember this construction, they are likely to end up *believing* his or her own lie. This is a likely explanation of Brian Williams' experience.

The low effort required to produce a denial has different consequences. Because there is no need to construct additional detail to persuade someone of its veracity, this means that there is little information about cognitive operations in memory. Because of this, liars are less likely to be able to remember that they had made a denial. However, the lack of elaboration also means that their memory is less likely to contain misleading perceptual and contextual detail. Because of this, it is less likely that deniers are going to have vivid false memories than confabulators.

The Impact of Lying on Memory

The SMF (Johnson et al., 1993) provides a number of predictions about liars' ability to remember both their lies and the original "event." These predictions have been explored in a relatively small number of studies.

Polage (2004, 2012) used a procedure based upon the *imagination inflation* paradigm (Garry, Manning, Loftus, & Sherman, 1996) to examine how lying about an unexperienced childhood event might influence memory for the event. In these experiments, participants were asked to create detailed false narratives about childhood events that had not happened to them. Polage (2004) found that after constructing a detailed, believable story about a childhood event that did not occur, the majority of participants rated their false narratives as less believable, and were more confident that the events did not happen to them. Presumably, the act of fabricating a rich narrative

account of an unexperienced event produced memories of the cognitive oper-
ations involved in generating those narratives, and participants used these
as cues that the memory had not been experienced. Subsequent research by
Polage (2012) found that when source monitoring ability was poor, either
at an individual difference level (Experiment 1) or due to a delay between
the lie generation session and the test session (Experiment 2), people were
more likely to exhibit "fabrication inflation" and believe their self-gener-
ated lies. Thus, it appears that false descriptions may increase false memories
about unexperienced events should people fail to remember generating those
descriptions.

Lane and colleagues attempted to evaluate the predictions of SMF in more
detail. Specifically, this research explored a broader set of variables (type of
lie and repetition), using a source memory test that allowed for a more pre-
cise characterization of memory for lies and truth-telling (Dianiska et al.,
2016; Vieira & Lane, 2013). In these studies, participants first saw a series
of objects (Vieira & Lane) or performed a series of actions (Dianiska et al.).
Subsequently, they were seated at a computer in front of a video camera. The
computer provided the name of the item or action and asked them to truth-
fully or deceptively describe the item (action) that they had seen (performed),
or deny that they had done so. Participants then either lied or told the truth
regarding these items on camera. An equal number of items (actions) were
talked about on camera once or three times (another set of items were seen
at encoding but were not discussed on camera, and subsequently appeared
on the test). Participant ratings after this phase revealed that they thought the
effort required for descriptions was significantly higher than for denials. After
a delay (48 hours in Vieira & Lane, 2013; 1 week in Dianiska et al., 2016),
participants were asked to indicate for each item whether they had seen it
(performed the action), whether they had lied or told the truth on camera,
and whether they had done so by describing or denying it.

Findings were consistent with predictions. Participants in all experiments
were substantially less likely to remember their denials than their descriptions.
Repetition also increased source accuracy. With respect to memory errors
(i.e., false memories of having seen pictures or performed an action), there
were generally two types of errors. First, *repeated* truthful denials increased
false memory. In other words, repeatedly denying that you had seen an item
you had not actually seen led you to subsequently falsely remember having
seen it. Furthermore, the results of the later experiments (Dianiska et al.)
were consistent with the idea that this error was based on the fact the item
had become more familiar and increased guessing. Second, we found that
participants falsely remembered seeing things they had confabulated (this
was particularly true at the longer delay). Additional data was consistent with
the notion that these errors were based more on false recollection of details.
Furthermore, repetition did not increase this type of error, and findings
suggested that the lack of an increase was because repetition increased the

likelihood that participants would remember generating the description and use this information to reject having seen it.

One matter that has been raised in the literature is whether denials can potentially impair memory for an original item. For instance, Otgaar and colleagues (Otgaar et al., 2014, 2016) examined the influence of false denials on later memory in both children and adults. When forced to falsely deny details present in a video (Otgaar et al., 2014) or present in studied pictures (Otgaar et al., 2016) during an initial interview, individuals were more likely to deny having *talked* to an experimenter about those details that were presented. However, their memory for the details themselves was not affected. Studies using our paradigm, which is ideal for detecting this type of impairment because we have control items—"studied" items that are not "rehearsed" (i.e., they do not appear in the second phase of the experiment, but are on the final test)—has found no evidence for such impairment. Across all three experiments (Dianiska et al., 2016; Vieira & Lane, 2013), denials never decrease memory below the level of never-rehearsed items. For instance, in Vieira and Lane (2013), once-denied items were recognized as having been studied 74% of the time (and 75% for thrice-denied items), while never-rehearsed items were recognized as having been studied 53% of the time. Thus, items that were denied actually show enhanced recognition performance.

Section Summary and Implications

As we have seen, the impact of lying on memory is well-predicted by the SMF (Johnson et al., 1993) and other basic theories of memory. The results are also consistent with other related areas of research examining false memory, including eyewitness suggestibility (e.g., Loftus, Miller, & Burns, 1978; Zaragoza & Lane, 1994) or the effects of fluency (e.g., Jacoby, Woloshyn, & Kelley, 1989). Despite these strengths, the literature on the topic is relatively small. To date, this research suggests that the type of lie one tells can influence what one remembers about an event and the act of lying. The constructive processes associated with creating an elaborated confabulation tend to increase the likelihood that the lie will be remembered relative to lying by a brief denial. Note that this does not mean that denials are always poorly remembered relative to confabulations. SMF predicts that the key issue is the constructive processing required by the lie. To the extent a denial requires more effortful processing (e.g., an elaborated claim of why someone could not have committed a crime), memory should also be enhanced. Results have also demonstrated that lying can also lead to false memories, and these false memories can arise from different mechanisms. For example, it appears repeated truthful denials may increase the fluency of those memories, and that confabulations appear more vividly remembered when people have forgotten the act of lying.

Despite the relatively small size of the literature, research on lying and memory already has potential implications for forensic contexts. Memory is required to maintain lies. During a criminal investigation, for instance, it is common for witnesses and suspects to be interviewed multiple times (Fisher, 1995). Furthermore, liars often maintain a strategy whereby they attempt to appear consistent (and thus not arouse suspicion) by trying to provide the same account across multiple statements (Fisher, Vrij, & Leins, 2013). Results suggest that interviewers may be able to take advantage of the fact that suspects are less likely to remember having denied having done something than if they produce a more elaborated characterization of their whereabouts. These studies also suggest potential concerns for the interview room, as innocent, truthful interviewees may begin to doubt the veracity of their own claims after repeated denials (see Henkel & Coffman, 2004, for a similar suggestion).

The Implications of Memory and Cognition for Detecting Deceit

Research on the ability to discriminate lies from truths has demonstrated that lie detection is a difficult, and not particularly accurate, task. When classifying truth-tellers and liars, people tend to hover around chance performance, showing a 53% accuracy rate across studies (Bond & DePaulo, 2006). Trained professionals in law enforcement demonstrate similar accuracy rates in detecting lies from truths, though with considerable overconfidence in their ability to do so (e.g., Meissner & Kassin, 2002). Part of the reason for this near-chance accuracy could be that the verbal and nonverbal cues that investigators are trained to use (and what lay people believe are reliable indicators) are not indicative of veracity. In 2003, DePaulo and colleagues examined 158 behavioral indicators of deception and found that only a minority of those cues were significantly, although weakly, related to deception. The cues that were most diagnostic of deception, such as level of detail and logical structure, were primarily cognitive and memory-based (DePaulo et al., 2003). Additionally, only a few nonverbal (Sporer & Schwandt, 2007) and paraverbal (Sporer & Schwandt, 2006) cues have been shown to be reliably associated with deception.

Popular training programs that advocate using nonverbal cues to improve deception detection may bias investigators to judge targets as overly deceitful (Meissner & Kassin, 2002; Kassin, Meissner, & Norwick, 2005). Further, while training itself has been shown to improve the ability to detect deception, a meta-analysis of available studies suggests that the most successful training programs focus on verbal content and memory-based cues (see Hauch, Sporer, Michael, & Meissner, 2016). Theoretical frameworks that consider memory and source monitoring have been used to further develop

approaches that focus on more diagnostic cues to deception, including criteria-based content analysis and reality monitoring.

Researchers have also begun to translate an understanding of the cognitive processes involved in deception into interviewing approaches that can be used to magnify differences between liars and truth-tellers. For example, interviewing techniques that manipulate the cognitive load of liars and truth-tellers, such as recalling a story in reverse order, or techniques that serve to enhance one's memory, such as the cognitive interview, can be employed to enhance indicators of deception (Vrij & Granhag, 2012; Fisher & Geiselman, 1992). In a recent meta-analysis, Vrij, Fisher, and Blank (2017) showed that compared to standard interviewing approaches, these cognitive lie detection approaches increase both truth and lie detection. In this section, we begin by reviewing the research literature on approaches to detect deception via the memorial content of a person's statement, followed by a discussion of interviewing approaches that manipulate cognitive load and memory-based cues.

Detection via Lie Content

Researchers have begun to focus on the content of statements that interviewees provide as a way of discriminating liars and truth-tellers. Based on the assumption that a statement derived from an actual experience (or memory) is characteristically different from a statement derived from an imagined (or invented) event (the Undeutsch hypothesis; Steller, 1989), two related sets of criteria have been developed to analyze the veracity of verbal statements. The first set of criteria were developed as part of the statement validity analysis tool to distinguish true and false allegations of sexual abuse by children (Steller & Köhnken, 1989). Criteria-based content analysis (CBCA) uses a present/absent judgment of interview statements across 19 criteria believed to be indicative of truth-tellers: thirteen of these criteria include indicators of veracity that would be difficult to fabricate for someone who did not genuinely experience the event, five are associated with a truthful person's motivation, and a final criterion is related to offense-specific information.

In general, CBCA assumes that true statements are more likely to be coherent, presented in non-chronological order, and contain greater quantities of detail than false statements. The types of detail specific to true statements include contextual embeddings (such as references to temporal or spatial details), descriptions of interactions, and reproductions of conversations. People being deceptive are more likely to be concerned with *appearing* truthful, so truthful statements are ironically more likely to contain information inconsistent with what an interviewer may believe as truthfulness. For example, a truthful statement is associated with spontaneous corrections and admitted lack of memory—both of which investigators more typically associate with deception. In a review of the literature, CBCA criteria were reliably more present in children's (Amado, Arce, & Fariña, 2015) and adult's

(Amado, Arce, Fariña, & Vilariño, 2016) truthful statements than in fabricated statements. In children, all CBCA criteria were significantly discriminative of truthful and fabricated accounts, with the largest effects from the criteria quantity of details and details characteristic of the offense. All but two CBCA criteria were predictive in adults, with the quantity of details criterion again associated with the largest effect size.

A second approach to identifying criteria was derived from research on how people attribute memories to external sources (i.e., experienced) versus internal sources (i.e., imagined). This process, termed *reality monitoring* (Johnson & Raye, 1981), can be used for both personal and interpersonal (i.e., judgments of whether *other people* have truly experienced a remembered event) reality monitoring decisions (e.g., Sporer & Sharman, 2006; Sporer, 2004). These judgments are based on evaluations of qualitative differences in the amount of sensory, perceptual, temporal, and spatial details present in a narrative. Memories of real experiences are likely to contain more perceptual and contextual details, given the presence of perceptual processes at encoding. In contrast, memories of imagined events that are internally derived are likely to contain more cognitive operations, such as thoughts and inferences made about the event. Empirical support for characteristics of internally derived events appears to be weak, whereas criteria associated with real experiences, such as contextual information and realism, are more discriminant of lies and truths (Masip, Sporer, Garrido, & Herrero, 2005; Memon, Fraser, Colwell, Odino, & Mastroberardino, 2010).

Recently, researchers have worked to integrate different sets of criteria to create a theoretically and empirically based credibility assessment tool. Comprised of cues derived from CBCA, reality monitoring, and other theories of deception, the assessment criteria indicative of deception (ACID; Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2007) and psychologically-based credibility assessment tool (PBCAT; Evans, Michael, Meissner, & Brandon, 2013; Evans & Michael, 2014) have been shown to guide users toward more reliable cues to deception and therein increase discrimination accuracy. In conjunction with an interview tactic that increases cognitive load and requires multiple recall attempts (discussed below), the ACID assessment tool can be used to compare the unique external, contextual, and internal details provided by interviewees during each phase of the interview. Training a variety of forensic professionals (Colwell et al., 2012) and police officers (Colwell, James-Kangal, Hiscock-Anisman, & Phelan, 2015) on the use of ACID has led to significant improvements in detection accuracy. Similarly, use of the PBCAT has been shown to increase accurate discrimination of true and false alibi statements, particularly when those statements are elicited by techniques that enhance cognitive load (Evans et al., 2013; Evans & Michael, 2014).

A final content cue to discriminate between lies and truths focuses on a liar's information dilemma—providing enough detail to appear truthful and

escape suspicion, but not too many details that might further an investiga-
tion. To overcome these competing motivations, liars appear to provide more
unverifiable details over details that could be verified. Nahari, Vrij, and Fisher
(2014) found that discrimination of liars and truth-tellers based upon the
number of verifiable details (versus unverifiable details, or a combination of
the two) led to better discrimination. The utility of verifiable details has also
been examined in the context of alibi witnesses (Nahari & Vrij, 2014) and
insurance claims (Harvey, Vrij, Nahari, & Ludwig, 2017; Nahari, Leal, Vrij,
& Warmelink, 2014).

Cognitive Load Techniques

Compared to truth-telling, lying is a more cognitively demanding task. In
addition to generating the lie with enough detail to be plausible and con-
vincing, a liar must simultaneously inhibit the automatically activated truth
response. Further, a liar will be burdened by monitoring his or her own
behavior as well as assessing the interviewer's behavior to ensure that they are
being believed (see Vrij, 2015, for a review). In this context, lying requires
considerable cognitive resources and therein takes more time than telling the
truth (Suchotzki et al., 2017). A truth teller, in contrast, is more likely to
believe in the power of his or her innocence (Kassin & Norwick, 2004) and
therein may be less concerned about appearing innocent.

   To the extent that lying *is* more cognitively demanding, considerable
research has shown that an interviewer can implement techniques during an
interview to exploit differences in the amount of cognitive load experienced
by interviewees and therein magnify the cues produced by liars (when com-
pared with truth-tellers; see Vrij et al., 2017). Techniques that have been
shown to improve discrimination include having interviewees recall a story
in reverse order (Evans et al., 2013; Vrij et al., 2008), instructing interview-
ees to maintain eye contact with an interviewer (Vrij, Mann, Leal, & Fisher,
2010), asking interviewees to complete a secondary task while telling their
stories (Debey, Verschuere, & Crombez, 2012; Visu-Petra, Varga, Miclea,
& Visu-Petra, 2013), or forcing two or more interviewees to "take turns"
when they are interviewed at the same time (Vernham, Vrij, Mann, Leal,
& Hillman, 2014). In addition to being assessed empirically, these tech-
niques have also been successfully taught to investigators (Vrij, Mann, Leal,
Vernham, & Vaughan, 2016).

   Evans et al. (2013) examined the cues present in liars and truth-tellers
when telling their stories in forward chronological order and reverse chron-
ological order (see also Vrij et al., 2008). Compared to those interviewed in
forward order, liars recalling in reverse order rated the experience as more
cognitively demanding. Further, the instruction to recall in reverse order
enhanced cues to deception associated with memory and cognition, such as
liars providing fewer auditory, spatial, and temporal details, providing less

plausible accounts, and appearing to struggle with recalling the event. Naive observers who were asked to identify liars and truth-tellers were 30% more accurate in the reverse order condition, particularly with respect to identifying liars.

When given the freedom to choose the content of a lie, liars may choose to report details from a previously experienced event, rather than fabricate new information (Leins, Fisher, & Ross, 2013). As a result, liars who have a pre-pared story prior to an interview may exhibit fewer cues to deceit than unprepared liars. One way in which interviewers can overcome a liar's preparation is to ask questions that the liar would not have anticipated. In comparing the responses of pairs of liars and pairs of truth-tellers, Vrij et al. (2009) found that participants' responses to unexpected questions, such as those regarding spatial layout (e.g., "where were you in relation to the other diners?") and requests to draw a sketch, enhanced discrimination between liars and truth-tellers who had previously had time to rehearse their stories.

Memory-Enhancing Techniques

Given the relationship between memory and certain content cues to deception discussed previously, researchers have also surmised that interview techniques that work to enhance memory will also magnify the key differences between liars and truth-tellers. Memory-enhancing techniques that have been assessed include the eye closure instruction (Perfect et al., 2008), mental reinstatement of context (Smith & Vela, 2001), sketch drawing (Leins, Fisher, Vrij, Leal, & Mann, 2011; Vrij et al., 2010), and providing subjects with a "model" statement about an unrelated topic (Leal, Vrij, Warmelink, Vernham, & Fisher, 2015).

One of the most effective memory-enhancing interview protocols developed in recent years is the cognitive interview (CI; Fisher & Geiselman, 1992). The CI comprises a set of retrieval-enhancing mnemonic techniques designed to increase the amount of information provided by a cooperative witness. The CI is characterized by a retrieval process wherein witnesses are encouraged to report everything that comes to mind during the recall process, but to withhold information that they are uncertain of (i.e., guessing). When combined with mnemonic approaches, context reinstatement, and eye closure, this approach leads to large gains in correct information recalled without a significant impact on accuracy (Memon, Meissner, & Fraser, 2010). The CI has been assessed in terms of information gain from suspects (e.g., Evans, Meissner, Ross, Houston, Russano, & Horgan, 2013) as well as distinguishing between true and false intentions (e.g., Sooniste, Granhag, Strömwall, & Vrij, 2015). Importantly, the CI has also been shown to improve credibility assessment. Köhnken, Schimossek, Aschermann, and Höfer (1995) compared true and false statements produced by a cognitive versus a structured interview with respect to a subset of CBCA criteria. While

the CI led to greater amounts of detail overall, it also improved discrimination of true and false statements based on CBCA criteria. Mnemonic elements of the CI, such as reverse order recall and drawing a sketch (described above), have also been shown to both enhance memory recall and improve discrimination of liars and truth-tellers, leading Geiselman (2012) to advocate for use of the CI with suspects.

Another technique found to increase differences between liars and truth-tellers involves providing subjects with an example statement about an unrelated topic (i.e., a "model statement"; Leal et al., 2015). Brief statements provide fewer cues to deceit, making it easier for liars to appear as though they are telling the truth; however, more cognitive content cues should be present when subjects are encouraged to provide more detailed statements. When truth-tellers encounter a model statement that offers substantial details, they adjust their expectations of what the interviewer considers a sufficient level of detail and accordingly provide more information (Bogaard, Meijer, & Vrij, 2014; Ewens et al., 2016; Leal et al., 2015). Alternatively, liars who are provided a model statement feel pressure to incorporate more false details into their narratives. Leal et al. (2015) found that after being exposed to a model statement, respondents provided statements that were more detailed; however, the details added by liars were less plausible than truth-tellers. The model statement has also been shown to increase the number of verifiable details provided by subjects, leading to a 25% increase in the classification accuracy of liars versus truth-tellers (Harvey, Vrij, Leal, Lafferty, & Nahari, 2017).

Forgetting and Statement Consistency

While cognitive load approaches and memory-based techniques are intended to make the task of lying more challenging and therein enhance the frequency of more diagnostic cues to deception, it is possible that certain conditions could diminish the frequency of memory-based cues to deception in truth-tellers' accounts. For example, when being questioned about an event that occurred long ago, truth-tellers may not have access to an event memory as a product of processes such as interference and decay—leading them to offer narratives that lack the most diagnostic cues to credibility discussed previously. Harvey, Vrij, Leal, Hope, and Mann (2017) observed that truth-tellers interviewed following a three-week delay disclosed fewer details about an event compared with those interviewed immediately after the event. Further, truth-tellers and liars did not differ with respect to the amount of detail reported following the delay. It is important, therein, that interviewers recognize that certain conditions that diminish memory quantity and quality, such as significant delays, can lead truth-tellers to appear more like liars.

Inconsistencies are often treated by both laypeople and professionals as an important indicator of deception; however, research suggests that the relationship between deception and inconsistency is more nuanced and may

depend upon the type of inconsistency present in a statement (Fisher et al., 2013). Inconsistencies in which information is either added during a subsequent interview (*reminiscence*) or omitted at a subsequent interview (*forgetting*) are more likely to be accurate than inconsistencies that are *contradictory* between two interviews. As such, consistency across statements or across respondents may be more indicative of liars who have rehearsed their statements. Asking unanticipated questions, as described previously, may offer a more diagnostic method for identifying liars. Pairs of liars and truth-tellers in Vrij et al. (2009) were interviewed with both anticipated questions and unanticipated questions after having time to prepare their answers together. Responses for anticipated questions were much more consistent across participants, likely due to the pair having had that time to rehearse their responses; however, responses to unanticipated questions were rather inconsistent across pairs of liars, providing a diagnostic cue to deception.

Ultimately, reliance upon consistency as an indicator of veracity can lead to misunderstandings within the criminal justice system. An innocent person who provides an initially mistaken alibi (due to faulty memory) may experience downstream direct and indirect effects on his or her likelihood of conviction (Crozier, Strange, & Luke, 2017). A mistaken alibi that is later corrected, for instance, will often be seen as inconsistent (and therein indicative of deception) by police officers and triers of fact. Such an inconsistency between statements may thus draw unwarranted suspicion on the alibi provider, potentially redirecting the course of an investigation toward an innocent suspect. Herein, it is important that investigators consider the type of inconsistency and the memory inherent to the recall attempt—to the extent that memory enhancing or strategic questioning strategies are applied, misattributions of deception (and guilt) can be avoided.

Section Summary

The science underlying deception and memory can inform efforts to increase lie detection accuracy. Research suggests that investigators should attend to the contents of a narrative and the cognitive operations evoked by the questioning. Further, inducing cognitive load and using techniques that enhance memory and reporting by the interviewee can enhance the most diagnostic cues to deception and therein improve judgments of credibility. Focusing on the elicitation of verifiable details also offers a promising approach to discriminating liars and truth-tellers, and furthering an investigation.

Conclusion

Brian Williams' odyssey (he has since returned to the newsroom) illustrates the perils of even small embellishments on memory for the truth. In his case, it is likely that his memory for the construction of his narrative faded,

while he retained a vivid memory that was the product of his imagination. Although such memories are problematic, law enforcement or intelligence professionals typically concern themselves with much more substantial deceit, such as when someone lies to avoid being charged with a capital crime. It is in this context that we see the many potential benefits of a deeper understanding of memory and deception. The ability to deceive others relies on the use of cognitive processes to construct a believable falsehood and maintain that lie over time. As we have seen, memories of lies are influenced by factors that include the type of lie that is told and the amount of times that a lie is repeated. In addition, the content of memories of actual and fabricated events differ in characteristic ways, and people can be trained to utilize these features to discriminate between them. Furthermore, it is possible to employ interviewing techniques to magnify differences in the reports of liars and truth-tellers to increase detection. Altogether, our review also illustrates another important point. Work in this domain clearly highlights the benefits of a research approach that takes as its foundation basic research and theory, while considering the constraints inherent in applied settings (Lane & Meissner, 2008). Understanding memory's role in deception has and will continue to motivate better and more nuanced ways of catching liars.

References

Amado, B. G., Arce, R., & Fariña, F. (2015). Undeutsch hypothesis and criteria based content analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context, 7*(1), 3–12. https://doi.org/10.1016/j.ejpal.2014.11.002.

Amado, B. G., Arce, R., Fariña, F., & Vilariño, M. (2016). Criteria-based content analysis (CBCA) reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology, 16*(2), 201–210. https://doi.org/10.1016/j.ijchp.2016.01.002.

Bogaard, G., Meijer, E. H., & Vrij, A. (2014). Using an example statement increases information but does not increase accuracy of CBCA, RM, and SCAN. *Journal of Investigative Psychology and Offender Profiling, 11*(2), 151–163. https://doi.org/10.1002/jip.1409.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Colwell, L. H., Colwell, K., Hiscock-Anisman, C. K., Hartwig, M., Cole, L., Werdin, K., et al. (2012). Teaching professionals to detect deception: The efficacy of a brief training workshop. *Journal of Forensic Psychology Practice, 12*(1), 68–80. https://doi.org/10.1080/15228932.2012.629592.

Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007). Assessment criteria indicative of deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling, 4*(3), 167–180. https://doi.org/10.1002/jip.73.

Colwell, K., James-Kangal, N., Hiscock-Anisman, C., & Phelan, V. (2015). Should police use ACID? Training and credibility assessment using transcripts versus

recordings. *Journal of Forensic Psychology Practice, 15*(3), 226–247. https://doi.org/10.1080/15228932.2015.1035187.

Crozier, W. E., Strange, D., & Loftus, E. F. (2017). Memory errors in alibi generation: How an alibi can turn against us. *Behavioral Sciences & the Law, 35*(1), 6–17. https://doi.org/10.1002/bsl.2273.

Debey, E., Verschuere, B., & Crombez, G. (2012). Lying and executive control: An experimental investigation using ego depletion and goal neglect. *Acta Psychologica, 140*(2), 133–141. https://doi.org/10.1016/j.actpsy.2012.03.004.

DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

Dianiska, R. E., Lane S. M., & Cash, D. K. (2016). *Type of lie differentially influences forgetting and false memory.* Paper presented at the meeting of the American Psychology-Law Society, Atlanta, GA.

Evans, J. R., Meissner, C. A., Ross, A. B., Houston, K. A., Russano, M. B., & Horgan, A. J. (2013). Obtaining guilty knowledge in human intelligence interrogations: Comparing accusatorial and information-gathering approaches with a novel experimental paradigm. *Journal of Applied Research in Memory and Cognition, 2*(2), 83–88. https://doi.org/10.1016/j.jarmac.2013.03.002.

Evans, J. R., & Michael, S. W. (2014). Detecting deception in non-native English speakers. *Applied Cognitive Psychology, 28*(2), 226–237. https://doi.org/10.1002/acp.2990.

Evans, J. R., Michael, S. W., Meissner, C. A., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a psychologically based credibility assessment tool. *Journal of Applied Research in Memory and Cognition, 2*(1), 33–41. https://doi.org/10.1016/j.jarmac.2013.02.002.

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., et al. (2016). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology, 30*(6), 854–862. https://doi.org/10.1002/acp.3270.

Fisher, R. P. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy, & Law, 1,* 732–764. https://doi.org/10.1037/1076-8971.1.4.732.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques in investigative interviewing: The cognitive interview.* Springfield, IL: Charles C Thomas.

Fisher, R. P., Vrij, A., & Leins, D. A. (2013). Does testimonial inconsistency indicate memory inaccuracy and deception? Beliefs, empirical research, and theory. *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment* (pp. 173–189). New York: Springer.

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review, 3*(2), 208–214. https://doi.org/10.3758/BF03212420.

Geiselman, R. E. (2012). The cognitive interview for suspects (CIS). *American Journal of Forensic Psychology, 30*, 5–20. Retrieved from http://www.thectkgroup.com/pdf/cognitiveinterview.pdf.

Harvey, A. C., Vrij, A., Leal, S., Hope, L., & Mann, S. (2017). Deception and decay: Verbal lie detection as a function of delay and encoding quality. *Journal of Applied Research in Memory and Cognition, 6*(3), 306–318. https://doi.org/10.1016/j.jarmac.2017.04.002.

Harvey, A. C., Vrij, A., Leal, S., Lafferty, M., & Nahari, G. (2017). Insurance based lie detection: Enhancing the verifiability approach with a model statement component. *Acta Psychologica, 174*, 1–8. https://doi.org/10.1016/j.actpsy.2017.01.001.

Harvey, A. C., Vrij, A., Nahari, G., & Ludwig, K. (2017). Applying the verifiability approach to insurance claims settings: Exploring the effect of the information protocol. *Legal and Criminological Psychology, 22*(1), 47–59. https://doi.org/10.1111/lcrp.12092.

Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2016). Does training improve the detection of deception? A meta-analysis. *Communication Research, 43*(3), 283–343. https://doi.org/10.1177/0093650214534974.

Henkel, L. A., & Coffman, K. J. (2004). Memory distortions in coerced false confessions: A source monitoring framework analysis. *Applied Cognitive Psychology, 18*, 567–588. http://dx.doi.org/10.1002/acp.1026.

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118,* 115–125. http://dx.doi.org/10.1037/0096-3445.118.2.115.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28. https://doi.org/10.1037/0033-2909.114.1.3.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67–85. https://doi.org/10.1037/0033-295X.88.1.67.

Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior, 29*(2), 211–227. https://doi.org/10.1007/s10979-005-2416-9.

Kassin, S. M., & Norwick, R. J. (2004). Why people waive their Miranda rights: The power of innocence. *Law and Human Behavior, 28*(2), 211–221. https://doi.org/10.1023/B:LAHU.0000022323.74584.f5.

Köhnken, G., Schimossek, E., Aschermann, E., & Hofer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology, 80,* 671–684. https://doi.org/10.1037/0021-9010.80.6.671.

Lane, S. M., & Meissner, C. A. (2008). A 'middle road' approach to bridging the basic-applied divide in eyewitness identification research. *Applied Cognitive Psychology, 22,* 779–787. https://doi.org/10.1002/acp.1482.

Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. P. (2015). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology, 20*(1), 129–146. https://doi.org/10.1111/lcrp.12017.

Leins, D., Fisher, R., & Ross, S. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology, 18*(1), 141–151. https://doi.org/10.1111/j.2044-8333.2011.02041.x.

Leins, D., Fisher, R. P., Vrij, A., Leal, S., & Mann, S. (2011). Using sketch drawing to induce inconsistency in liars. *Legal and Criminological Psychology, 16*(2), 253–265. https://doi.org/10.1348/135532510X501775.

Lindsay, D. S. (2008). Source monitoring. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory. Vol. 2 of learning and memory: A comprehensive reference*, 4 vols (J. Byrne, Editor) (pp. 325–348). Oxford: Elsevier.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(1), 19–31. https://doi.org/10.1037/0278-7393.4.1.19.

Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law, 11*(1), 99–122. https://doi.org/10.1080/106831604 10001726356.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*(5), 469–480. https://doi.org/10.1023/A:1020278620751.

Meissner, C. A., Kelly, C. E., & Woestehoff, S. A. (2015). Improving the effectiveness of suspect interrogations. *Annual Review of Law and Social Science, 11,* 211–233. https://doi.org/10.1146/annurev-lawsocsci-120814-121657.

Memon, A., Fraser, J., Colwell, K., Odinot, G., & Mastroberardino, S. (2010). Distinguishing truthful from invented accounts using reality monitoring criteria. *Legal and Criminological Psychology, 15*(2), 177–194. https://doi.org/10.1348/135532508x401382.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16,* 340–372. https://doi.org/10.1037/a0020518.

Nahari, G., Leal, S., Vrij, A., Warmelink, L., & Vernham, Z. (2014). Did somebody see it? Applying the verifiability approach to insurance claim interviews. *Journal of Investigative Psychology and Offender Profiling, 11*(3), 237-243. https://doi.org/10.1002/jip.1417.

Nahari, G., & Vrij, A. (2014). Can I borrow your alibi? The applicability of the verifiability approach to the case of an alibi witness. *Journal of Applied Research in Memory and Cognition, 3*(2), 89–94. https://doi.org/10.1016/j.jarmac.2014.04.005.

Nahari, G., Vrij, A., & Fisher, R. P. (2014). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology, 19*(2), 227–239. https://doi.org/10.1111/j.2044-8333.2012.02069.x.

Otgaar, H., Howe, M. L., Memon, A., & Wang, J. (2014). The development of differential mnemonic effects of false denials and forced confabulations. *Behavioral Sciences & the Law, 32*(6), 718–731. https://doi.org/10.1002/bsl.2148.

Otgaar, H., Howe, M. L., Smeets, T., & Wang, J. (2016). Denial-induced forgetting: False denials undermine memory, but external denials undermine belief. *Journal of Applied Research in Memory and Cognition, 5*(2), 168–175. https://doi.org/10.1016/j.jarmac.2016.04.002.

Perfect, T. J., Wagstaff, G. F., Moore, D., Andrews, B., Cleveland, V., Newcombe, S., et al. (2008). How can we help witnesses to remember more? It's an (eyes) open and shut case. *Law and Human Behavior, 32,* 314–324. https://doi.org/10.1007/s10979-007-9109-5.

Polage, D. C. (2004). Fabrication deflation? The mixed effects of lying on memory. *Applied Cognitive Psychology, 18*(4), 455–465. https://doi.org/10.1002/acp.995.

Polage, D. C. (2012). Fabrication inflation increases as source monitoring ability decreases. *Acta Psychologica, 139*(2), 335–342. https://doi.org/10.1016/j.actpsy.2011.12.007.

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*(2), 203–220. https://doi.org/10.3758/BF03196157.

Sooniste, T., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2015). Statements about true and false intentions: Using the cognitive interview to magnify the differences. *Scandinavian Journal of Psychology, 56,* 371–378. https://doi.org/10.1111/sjop.12216.

Sporer, S. L. (2004). Reality monitoring and the detection of deception. In P. A. Granhag & L. A. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 64–102). New York, NY: Cambridge University Press. https://doi.org/10.1017/cbo9780511490071.004.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20,* 421–446. https://doi.org/10.1002/acp.1190.

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13,* 1–34. https://doi.org/10.1037/1076-8971.13.1.1.

Sporer, S. L., & Sharman, S. J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology, 20*(6), 837–854. https://doi.org/10.1002/acp.1234.

Steel, E., & Somaiya, R. (2015, February 10). Brian Williams suspended from NBC for 6 months without pay. *New York Times.* www.nytimes.com/2015/02/11/business/media/brian-williams-suspended-by-nbc-news-for-six-months.html.

Steller, M. (1989). Recent developments in statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 135–154). Deventer, Holland: Kluwer. https://doi.org/10.1007/978-94-015-7856-1_8.

Steller, M., & Köhnken, G. (1989). Criteria-based statement analysis. Credibility assessment of children's testimonies in sexual abuse cases. In D. Raskin (Ed.), *Psychological techniques in law enforcement*, 217–245. New York: Springer-Verlag.

Suchotzki, K., Verschuere, B., Van Bockstaele, B., Ben-Shakhar, G., & Crombez, G. (2017). Lying takes time: A meta-analysis on reaction time measures of deception. *Psychological Bulletin, 143*(4), 428–453. https://doi.org/10.1037/bul0000087.

Vernham, Z., Vrij, A., Mann, S., Leal, S., & Hillman, J. (2014). Collective interviewing: Eliciting cues to deceit using a turn-taking approach. *Psychology, Public Policy and Law, 20,* 309–324. https://doi.org/10.1037/law0000015.

Vieira, K. M., & Lane, S. M. (2013). How you lie affects what you remember. *Journal of Applied Research in Memory and Cognition, 2*(3), 173–178. https://doi.org/10.1016/j.jarmac.2013.05.005.

Visu-Petra, G., Varga, M., Miclea, M., & Visu-Petra, L. (2013). When interference helps: Increasing executive load to facilitate deception detection in the concealed information test. *Frontiers in Psychology, 4.* https://doi.org/10.3389/fpsyg.2013.00146.

Vrij, A. (2015). Verbal lie detection tools: Statement validity analysis, reality monitoring and scientific content analysis. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches* (pp. 3–35). Chichester: Wiley.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22*(1), 1–21. https://doi.org/10.1111/lcrp.12088.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*(2), 110–117. https://doi.org/10.1016/j.jarmac.2012.02.004.

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., et al. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior, 33*(2), 159–166. https://doi.org/10.1007/s10979-008-9143-y.

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32*(3), 253–265. https://doi.org/10.1007/s10979-007-9103-y.

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010). 'Look into my eyes': Can an instruction to maintain eye contact facilitate lie detection? *Psychology, Crime & Law, 16*(4), 327–348. https://doi.org/10.1080/10683160902740633.

Vrij, A., Mann, S., Leal, S., Vernham, Z., & Vaughan, M. (2016). Train the trainers: A first step towards a science-based cognitive lie detection training workshop delivered by a practitioner. *Journal of Investigative Psychology and Offender Profiling, 13*(2), 110–130. https://doi.org/10.1002/jip.1443.

Zaragoza, M. S., & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 934–945. https://doi.org/10.1037/0278-7393.20.4.934.

Memory Detection: Past, Present, and Future

*Linda Marjoleine Geven, Gershon Ben-Shakhar,
Merel Kindt and Bruno Verschuere*

It was November 1985 when the bodies of two sexually assaulted, murdered, and mutilated children, a seven-year-old girl and her eight-year-old brother, were found in Plainfield, New Jersey. Byron Halsey, the boyfriend of the children's mother, quickly became the main suspect in the high-profile investigation. Was Mr. Halsey the true perpetrator of this crime? And could a lie detector help the case?

A Short History on Lie Detection

Most people are familiar with the Pinocchio effect; the nose of the wooden puppet instantaneously and observably grew whenever he told a lie. This famous story does not stand alone in the history of detecting deception. In ancient India, suspects were asked to chew raw rice and those who could not spit out the rice were in big trouble. It was believed that liars have a dry

L. M. Geven (✉) · M. Kindt · B. Verschuere
Department of Clinical Psychology, University of Amsterdam,
Amsterdam, The Netherlands
e-mail: L.M.Geven@uva.nl

M. Kindt
e-mail: M.Kindt@uva.nl

B. Verschuere
e-mail: B.J.Verschuere@uva.nl

L. M. Geven · G. Ben-Shakhar
Department of Psychology, Hebrew University of Jerusalem, Jerusalem, Israel

mouth, therefore making the rice stick to the tongue. Similar methods were used by the Bedouins of Arabia, using an even more brutal method. If the tongue of the suspect would stick to a burning hot iron, deception was indicated (Lykken, 1998; Trovillo, 1939).

Whether it is a growing nose or a dry mouth, no single bodily response has been established to be uniquely related to lying (Vrij, 2008). As a result, many Indians and Bedouins might have been wrongfully convicted. The problem lies in the flawed theory underlying the idea to use a dry mouth as a cue to deception. What seems to be forgotten is the fact that even an innocent person fearing to be falsely accused could experience increased stress and enhanced bodily responses when facing a lie detection test. Some contemporary deception detection methods, such as the Control Question Polygraph Test (CQT; Reid, 1945), are still based on the idea that fear or stress responses reveal deception. Since an incorrect outcome may put an entire investigation on the wrong track (Kassin, Bogart, & Kerner, 2012), invalid lie detection tests based on stress-induced cues should be avoided.

Byron Halsey, suspected of the molestation and brutal murder of the two children, was convicted to two life sentences plus 20 years, after failing the polygraph examination. Importantly, on key crime details such as the location of the bodies and the modus operandi, Halsey initially gave an incorrect narrative before guessing the correct manner of death and confessing to the crimes in a tense interrogative setting. More than two decades later, after spending most of his prison time in solitary confinement for his own safety, post-conviction DNA testing not only proved Halsey's innocence, but also implicated the true perpetrator of the horrific crime.

According to William Blackstone in his book *Commentaries on the Laws of England* (1830), exonerating ten guilty individuals is deemed better than wrongfully incarcerating a single innocent person. Since then, preventing false positive errors (i.e., mistakenly identifying an innocent person to be guilty) forms the basis of our legal system in which the scales of justice are tilted in favor of the accused unless sufficiently proven guilty. Deception detection methods should therefore ideally reach not only good sensitivity (i.e., proportion of correctly detected guilty suspects based on the test outcome), but particularly high levels of specificity (i.e., proportion of correctly identifying the innocents), in order that errors as the one made in the case of Byron Halsey are avoided. In high-stake situations, such as in the case of criminal proceedings, accurate and reliable deception detection techniques are an absolute necessity. Methods to detect deceit should be based on a sound scientific framework as to reliably indicate possible involvement in a crime and avoid wrongful incarceration.

MEMORY DETECTION

The fundamental difference between flawed polygraph tests fixated on deception-induced stress (Ben-Shakhar, 1991; Ben-Shakhar, Bar-Hillel, & Lieblich, 1986; Lykken, 1991) and methods designed to detect memory traces is that

the latter methods focus on detecting recognition of intimate crime details rather than deception. While lie detection attempts to determine deception by interpreting answers to interrogational questions such as "Did you kill the two children?" the purpose of the Concealed Information Test (CIT; first introduced as the Guilty Knowledge Test by Lykken, 1959, 1960) is to verify whether the suspect is aware of certain crime-related information, for instance, whether the murder weapon was a bomb, a firearm, or a knife. This method is therefore labeled a memory detection test rather than a lie detection test.

The objective of the CIT is to verify whether the suspect possesses crime-related information that only the perpetrator would be aware of. The method requires that the examiner determines a number of established facts from the investigation which only the true culprit will be able to recognize. Then, the examiner creates a CIT resembling a multiple-choice test with several questions, such that each question is comprised of one detail of the crime in question, and several neutral control items. In the case of a homicide, for instance, the CIT might involve questions concerning the murder weapon and the location of the victim in the crime scene. For example, "How was the victim murdered? (a) by beating, (b) by stabbing, (c) by drowning, (d) by shooting, or (e) by poisoning" or "Where was the victim attacked? (a) bathroom, (b) kitchen, (c) bedroom, (d) garden, (e) living room." For each question, there is only one item that reflects the correct feature of the crime under investigation (i.e., the critical or crime-related item, called the probe). The other options are neutral control items from the same category as the relevant item (i.e., called the irrelevants). These irrelevant items are chosen carefully, such that all options would seem equally plausible to unknowledgeable individuals. As a result, to innocent suspects who are unaware of the crime's details, all items will trigger similar responses. On the other hand, guilty suspects tend to react significantly different to the recognized crime-related detail than to the irrelevant items. This differential response (i.e., probe minus irrelevant response, labeled the CIT effect) indicates critical knowledge of the crime in question, which should lead to further investigation of the suspect (Lykken, 1974, 1998).

Imagine the case of the double-child murder in New Jersey. If Byron Halsey would have been guilty, it is likely that he remembered the location of the victims' bodies. Therefore, he would have shown a differential response to the correct alternative (i.e., basement). On the other hand, since the suspect was in fact innocent and therefore did not know and could not infer which alternative reflected the true feature of the crime, he would show similar responses to all items. By using several CIT questions, each with 4 or 5 alternatives and several dependent measures, the probability of a false positive outcome could be controlled by the investigators. The main advantage of the CIT over methods focusing on deception is its use of proper controls. The CIT establishes a within-person control in which responses to the critical alternative are compared to an estimate of the response to the correct alternative if the person would be in fact innocent (i.e., the irrelevant options).

Moreover, stress-induced mental states driven by the potential consequences of failing the test are expected to influence both relevant and irrelevant alternatives similarly. Thus, whether the suspect is calm, aroused, or even frightened, it is still expected that the response to the critical crime-related item is stronger than to the alternatives whenever the suspect recognizes the correct answer. Likewise, if the suspect is in fact innocent and unaware of the crime-related items, neither the present emotional state nor the possible consequences of failing the test can influence the CIT outcome, since it affects responses to both critical and control items. In addition, the CIT is a scientific approach to deception detection, substantiated by extensive theoretical and applied research (e.g., Ben-Shakhar, 2012; Verschuere, Ben-Shakhar, & Meijer, 2011).

Underlying Theory

The orienting response (OR) has long been the dominant hypothesis for the CIT effect (see Ben-Shakhar, 1977; Lieblich, Kugelmass, & Ben-Shakhar, 1970; Lykken, 1974). A combination of physiological and behavioral responses in reaction to an external novel stimulus forms the basis of the OR. Already in 1927, while working on his famous classical conditioning experiments, Pavlov noted that the dogs in his laboratory shifted their attention to new visitors. But it did not take long before their interest in this novel person would decline, resulting in a decrease in the dogs' investigatory response (Sokolov, 1963). This incidental finding demonstrated that the initial OR may gradually habituate in magnitude. However, when the stimulus holds a special significance to the subject, an enhanced OR can be observed (Sokolov, 1963). Significant stimuli have also been proven to be more resistant to habituation. Changes in the magnitude of the OR therefore allow for differentiation between salient and neutral stimuli (Bradley, 2009; Gamer, 2011; Lykken, 1974). This effect also forms the basis of the CIT. The critical crime detail (e.g., strangulation of the victim) holds a very significant meaning to the guilty individual, but not to the innocent suspect. Therefore, a consistently stronger response to "strangulation" compared to the control items (e.g., "drowning", "poisoning", etc.) is an indication that the suspect has knowledge about the crime in question and should be further investigated.

More recently, response inhibition was found to also underlie the CIT effect (klein Selle, Verschuere, Kindt, Meijer, & Ben-Shakhar, 2016, 2017; Suchotzki, Verschuere, Peth, Crombez, & Gamer, 2015; Verschuere, Crombez, Koster, Van Bockstaele, & De Clercq, 2007). Since it is assumed that the truthful answer is the natural and default response to a question, lying requires a significant amount of cognitive resources in order to actively suppress the truth (Suchotzki et al., 2015). Response inhibition is thus required to prevent the truth from being exposed in the CIT and to give a deceptive response instead. For the CIT rationale, it is reasoned that in order

to remain undetected, a guilty suspect must suppress the increased arousal associated with recognition of the critical item. However, this effort has the paradoxal consequence (Pennebaker & Chew, 1985) that it further increases physiological responses to the critical items.

Response Measures

Recognition of critical items in the CIT can be assessed by autonomic and behavioral measures, as well as brain-related measures, such as functional magnetic resonance imaging (fMRI) or electroencephalography (EEG). For all measures, the key factor in memory detection is the differential response to the critical items compared to the irrelevant options as an indicator of recognition. Autonomic nervous system (ANS) measures such as skin conductance, heart rate (HR), and respiration have been used since the beginning of memory detection. While measuring amplitudes of the galvanic skin response elicited by items in the CIT, larger skin conductance responses (SCR) upon probe presentation were found for individuals attempting to conceal information (Lykken, 1960). Moreover, respiration (RLL), measured with pneumatic straps around the chest and abdomen, is smaller upon recognition of the relevant items compared to neutral items. Similarly, cardiovascular measures can indicate concealed information. Phasic HR, measured with electrodes on the chest or with infrared at the fingertip, decreases within 15 seconds after presentation of the critical item, compared to irrelevant alternatives.

Does It Work?

Since the early 1960s, there is ample evidence for successfully detecting crime-related knowledge and discriminating between guilty and innocent individuals with the CIT. An early laboratory study on the validity of the CIT was conducted in 1959 by David Lykken, laying a promising foundation for future research. In an attempt to mimic real-life situations in which memory detection tests could be meaningful, a mock-crime paradigm was used. By measuring and ranking the amplitude of the galvanic skin response upon presentation of the probe and irrelevant alternatives for both guilty and innocent participants, around 90% of the participants were classified correctly. In a meta-analysis (Ben-Shakhar & Elaad, 2003), the validity of 42 subsequently executed mock-crime experiments revealed a very large average effect size (Cohen's d) of 2.09 (0.80 is considered a large effect size by Cohen, 1988). These results confirmed that the SCR measure can accurately detect relevant information and differentiate between individuals with and without knowledge of the critical mock-crime details.

In addition to skin conductance, various other experiments added different measures of the ANS such as changes in respiration and HR. In most observations, the SCR outperformed other measures in detecting recognition

(see, e.g., Balloun & Holmes, 1979; Podlesny & Raskin, 1978). While the SCR remains the most valid single autonomic measure, an accumulation of all three was found to provide incremental evidence for the effectiveness of the CIT (Gamer, Verschuere, Crombez, & Vossel, 2008). More recently, meta-analytic results reconfirmed the validity of these psychophysiological measures to detect the presence or absence of crime-related knowledge in a suspects' memory (Meijer, klein Selle, Elber, & Ben-Shakhar, 2014). Again, large effect sizes were found for the SCR, RLL, and HR (1.55, 1.11, and 0.89 respectively).

Another shift of interest took place in the last two decades toward the potential of the response latency measure for detecting deception and concealed knowledge. Relying on reaction times (RTs) to distinguish between innocents and individuals withholding critical information requires a slightly different procedure, yet the response difference between the probe and several irrelevant options remains essential. Initial research using handheld stopwatches did not find strong effects for RTs as an index of deception (see Luria, 1932; Marston, 1920), but computerized methods have led to a renewed research interest in the use of response latency to detect deception.

While measuring RTs, participants in a laboratory study (Seymour, Seifert, Shafto, & Mosmann, 2000) had to indicate whether they recognized the stimuli presented in the CIT by pressing one of two response keys. Critical details from the committed mock-crime were intermixed with neutral items. Upon measuring response latencies for denying knowledge of the probe in comparison with irrelevant words, 23 out of 27 participants were correctly classified as guilty. A recent meta-analysis based on studies relying on computerized RT measures showed the potential of the RT-CIT (Cohen's *d* of 1.30; Suchotzki, Verschuere, Van Bockstaele, Ben-Shakhar, & Crombez, 2017).

Interestingly, recent insights suggest that different response measures may be driven by different mechanisms. Specifically, it has been suggested that some measures (e.g., elevated skin conductance) may be mostly related to the concealed item drawing attention (i.e., OR), whereas other measures (e.g., the drop in HR, more shallow breathing cycles, and slowing of reaction time) reveal the subsequent deliberate concealment of the recognition (i.e., response inhibition; klein Selle et al., 2016, 2017; Rosenfeld, Oszan, & Ward, 2017; Suchotzki et al., 2015).

Whereas laboratory research allows for a controlled environment and manipulation of isolated variables, questions can be asked about the generalizability of the results to real-life cases (i.e., ecological validity). Establishing accuracy rates for memory detection in a realistic situation, such as criminal investigations, requires validity studies conducted in authentic settings. Therefore, the few field studies that were reported may provide an additional insight regarding the external validity of laboratory experiments designed to assess the validity of the CIT. These studies conducted with real suspects, for whom the stakes are high and who are motivated to avoid detection,

produced mixed results. Specifically, Elaad (1990) found a very high-accuracy rate (98%) for discerning innocent suspects with the SCR measure, but the rate for correctly identifying guilty suspects was much lower than expected (42%). In a second field study (Elaad, Ginton, & Jungman, 1992), measuring RLL in addition to SCR, the results were a bit more promising, although still far from optimal. Both measures separately could detect innocent and guilty suspects with 97% and 53% accuracy, respectively, yet combining SCR with RLL led to increased detection accuracy (76%) for guilty suspects. However, it is worth noting that the CITs applied in these field studies were not optimal. First, the authors used a scoring procedure that is nowadays replaced by improved computational systems in which multiple measures can be combined and standardized. Second, the number of questions was much smaller than recommended, and third, the CITs were administered immediately after a CQT, which might have attenuated the sensitivity of the measures due to habituation effects. Moreover, when conducting field studies, there might be other difficulties, mostly concerning the establishment of ground truth. In actual cases, it is very difficult to establish proof of whether the test outcome (either guilty or innocent) was in fact correct. Therefore, most field research uses confessions as the principle proof of actual guilt. Since this criterion is vulnerable to sampling biases (Iacono, 1991; Patrick & Iacono, 1991), the data should be considered with caution.

External Validity of CIT Studies

Due to the difficulties of establishing a solid ground truth criterion in realistic settings, several researchers have adopted an alternative approach for evaluating the external validity of CIT studies conducted in artificial laboratory conditions (for a review see Ben-Shakhar & Nahari, 2018). Specifically, researchers have identified several factors that differentiate between laboratory and realistic environments and manipulated each factor in controlled experiments. In the following sections, we briefly review this research focusing on each of the identified factors: levels of arousal, motivation to avoid detection, and the influence of a delay between crime and CIT, as well as effects of real-life deception on external validity.

Level of Arousal

Clearly, suspects undergoing criminal investigations are much more aroused than subjects participating in laboratory experiments. Indeed, the average HR of examinees in real-life tests is much higher than that of laboratory examinees (Verschuere, Meijer, & De Clercq, 2011). Early CIT studies relied on the card test paradigm, where subjects pick a card from a deck, hide this information, and a subsequent CIT is conducted to detect the hidden information. Evidently, this is very different from a realistic scenario in the legal field.

More recent studies adopted the mock-crime paradigm, which seems to better approximate realistic crimes. In their meta-analysis, Ben-Shakhar and Elaad (2003) compared the CIT effect with the SCR measure obtained from card test and mock-crime experiments. They found a much larger CIT effect for mock-crime than for card test studies (2.09 vs. 1.35, respectively). However, the level of arousal experienced by mock-crime participants is still quite moderate and far below what real suspects may experience during realistic polygraph tests. To better tackle this question, Verschuere et al. (2011) conducted a card test study with suspects undergoing realistic police polygraph interrogations. The enhanced arousal level in this condition was confirmed by a higher baseline HR than typically measured with research participants in the laboratory. Even under these higher levels of arousal, the CIT effect was found to exist in the field: HR, RLL, and SCR significantly changed upon presentation of the picked card as opposed to irrelevant cards. In a direct comparison of the CIT effect obtained in card tests conducted during a realistic polygraph investigation and laboratory experiment, Zaitsu (2016) reported similar effects in both settings. Additional studies attempted to examine whether the level of arousal affects the outcomes of the CIT in controlled experiments. These studies, which employed different types of arousal manipulation, reported similar CIT effects in the high- and low-arousal conditions (Bradley & Janisse, 1981; klein Selle, Verschuere, Kindt, Meijer, Nahari, & Ben-Shakhar, 2017; Kugelmass & Lieblich, 1966; Peth, Vossel, & Gamer, 2012).

Motivation to Avoid Detection

Another difference between actual examinations and laboratory simulations might be the motivation to avoid detection and appear innocent. While a guilty suspect might have sufficient reasons to keep up appearances, research participants obviously do not face comparable detrimental consequences. Motivational manipulations are commonly achieved by instructions (e.g., Gustafson & Orne, 1963), incentivizing participants for beating the polygraph (Bradley & Warfield, 1984), or punishing participants for an undesirable outcome (Lykken, 1974). Although experiments did not always reveal consistent findings (Gustafson & Orne, 1963, 1965; Horvath, 1978, 1979; Lieblich, Naftali, Shmueli, & Kugelmass, 1974), meta-analytic results support the notion that the SCR effect size increases when the motivation to avoid detection is high (Ben-Shakhar & Elaad, 2003; Meijer et al., 2014). Guilty suspects increasing their effort to deceive the test might therefore—paradoxically—show enhanced responses to the probe among irrelevant items and thus aid their own detection (e.g., motivational impairment hypothesis; DePaulo & Kirkendol, 1989). In contrast, CIT studies using response latency as the dependent variable do not seem to benefit from additional motivational instructions. In a meta-analytic study, liars under motivation instructions

to appear innocent were detected equally adequate as a control condition (Suchotzki et al., 2017).

Delay Between Crime and CIT

While in the typical CIT experiment the test is administered immediately after participants were exposed to the critical items, realistic polygraph tests are often administered several weeks, or even months after the crime (Ben-Shakhar & Furedy, 1990). Naturally, memory declines with time, and since the CIT is a memory detection test, this might pose as a pitfall in criminal investigations. Indeed, research findings confirm the weakening effect of a time delay on detection efficiency of the CIT (Carmel, Dayan, Naveh, Raveh, & Ben-Shakhar, 2003; Gamer, Kosiol, & Vossel, 2010; Nahari & Ben-Shakhar, 2011; Peth et al., 2012). However, this effect is mediated by the type of items used in the test. Whereas the memory for less-important, peripheral items decays quite rapidly, detection of central items or the gist of an event (i.e., items that are directly associated with the crime, such as the weapon that was used) typically persists. In a real-life scenario, more reliable responses can be expected to a question about the murder weapon than to a question regarding the clothes of the victim. Examiners are therefore advised to use central crime details that are likely to be better encoded and more easily recalled.

The Free Choice to Commit a Crime, Deceive, and Conceal Information

Another important distinction between laboratory experiments and realistic criminal investigations is the deliberate aspect of criminal actions. Deception is commonly defined as a voluntary act (see Vrij, 2004), in which intention is a key factor. Yet, in laboratory studies on detecting deception, participants are often explicitly instructed to commit a staged crime and subsequently conceal knowledge (e.g., Lykken, 1959; Nahari & Ben-Shakhar, 2011). More recently, researchers have begun to explore the role of instructed versus spontaneous cheating and lying (Blakemore, Winston, & Frith, 2004; Kozel et al., 2005; Mohamed et al., 2006; Sip, Roepstorff, McGregor, & Frith, 2008). For instance, Nahari, Breska, Elber, klein Selle, and Ben-Shakhar (2017) gave participants a free "choice" to decide whether to enact a mock-crime or an innocent computerized task and compared those who choose to commit the mock-crime with participants who were instructed to do so. The study revealed a similar CIT detection efficiency, based on SCR, RLL, and RT measures, in these two conditions.

However, deception, in all its complexity, can only be fully investigated when the decision to deceive is based entirely on the participants' own initiative. In an externally more valid paradigm, participants engaged in a trivia

quiz and were provided with a monetary incentive for high-accuracy performance. Participants were randomly allocated to either a condition where they were instructed to cheat on the quiz (mimicking the typical laboratory setup) or to a condition in which they were provided with the opportunity to cheat using Google, yet without explicit instructions to do so. Assessments of their RTs (Geven, Ben-Shakhar, Kindt, & Verschuere, 2018) and physiological responses (Geven, klein Selle, Ben-Shakhar, Kindt, & Verschuere, 2018) in the CIT revealed that both instructed and self-initiated cheaters showed a similar pattern upon recognition. The results indicate that the cognitive signature of lying is not restricted to explicitly instructed deception, but can also be observed for its real-life equivalent. These findings are highly encouraging from an ecological validity perspective, suggesting that when it comes to free choice and voluntary deception, the results of laboratory studies are a realistic reflection of the field.

LIMITATIONS OF THE CIT

In the previous sections, we emphasized the strength of the CIT as a scientific approach to memory detection, based on proper control questions with validity estimates generated from extensive laboratory research. We have also argued, when systematically examining several factors differentiating between the laboratory and realistic settings, that results of laboratory studies can be generalized. However, the CIT is by no means a perfect method, and it is important to discuss its limitations. There are two main factors that might limit the validity of the CIT in realistic applications: the potential effects of countermeasures and the danger that critical crime-related items may contaminate innocent suspects.

Countermeasures

While increased motivation might enhance the CIT effect in the field, fear of detection can also tempt guilty suspects to use countermeasures (i.e., consciously alter physiological reactions during the CIT to avoid detection). Two strategies can be applied to diminish the expected probe-irrelevant difference during the polygraph examination. Suspects can either try to suppress their responses to the relevant crime details or artificially enhance responses to neutral, irrelevant items (Ben-Shakhar, 2011) by using mental (e.g., demanding cognitive activities such as counting backward from 100 in steps of seven) or physical countermeasures (e.g., biting their tongue or moving their toes). Various experiments were conducted to assess the effects of these countermeasures on the outcomes of the CIT (e.g., Ben-Shakhar & Dolev, 1996; Elaad & Ben-Shakhar, 1991, 2009; Honts, Devitt, Winbush, & Kircher, 1996; Lykken, 1960; Peth, Suchotzki, & Gamer, 2016). These studies revealed that ANS measures are affected by both mental and physical

countermeasures and could thereby enhance the false negative rates, if suspects are aware of how to use countermeasures to their advantage. Yet, no increase in false positives is expected to occur.

Leakage of Crime-Related Information

Influenced by the use of mass media channels, news spreads easier than ever. Disclosure of information cannot always be prevented, yet this can alter the validity of memory detection test results. Besides news reports, crime-related information can unintentionally be leaked to suspects during their interrogation. For example, Byron Halsey came to know the location where the bodies were discovered after some guess-work. If the location of the victims would then have been used as a critical detail in the CIT, it might have triggered a false positive outcome, since the CIT effect is driven by recognition instead of actual guilt. Several studies have examined the effects of information leakage on the outcome of the CIT, and although their results are not entirely consistent, it seems that leakage of information to innocent suspects may significantly increase the rate of false positives (for a review, see Bradley, Barefoot, & Arsenault, 2011; Osugi, 2018). Thus, leakage remains a major obstacle, and it can be avoided only by adopting careful police investigation practices.

In addition to avoiding leakage, using multiple relevant items, and increasing the number of questions in a CIT, more research is needed to evaluate the validity of a more specific CIT. For example, instead of the known cause of death by "strangulation", it might be possible to ask a more specific question, such as, "What object was used to strangulate the victim? (a) rope, (b) shoe lace, (c) hands, (d) baton, or (e) wire." This may limit the risk of information leakage as only individuals with specific knowledge of the crime in question are likely to be exposed to this information.

Agenda for Future Research

Asking the Right Questions

In order to detect a memory trace that could link a suspect to the crime, it is important to ask the proper questions in the CIT. But can experienced examiners select the appropriate crime details? Rationally, central crime items are better remembered than peripheral items and are therefore also better detectable. Research revealed higher differential SCRs for central items after a delay (Gamer et al., 2010; Nahari & Ben-Shakhar, 2011). Besides centrality, actual enactment of the crime might result in more stable memory traces (Madan & Singhal, 2012) and allow for a more accurate distinction between innocent and guilty suspects in the CIT. This effect may be moderated by the saliency of the items. Since the OR is the bodily reaction upon presentation of significant stimuli, we tend to have a larger OR to, for example, a picture

of the victim than to a random stranger. Also in the RT-CIT, larger effects were found when stimuli that draw more attention were used (Suchotzki et al., 2017). Unfortunately, in real life, the control over how a complex crime scene is perceived is beyond the control of the examiners. It cannot be assumed that all details derived from the criminal investigation are actually noticed and stored in memory, ready to be exposed in a memory detection test. Future studies could therefore focus on the link between a perpetrators' memory and the stimuli to be tested.

Information Gathering Using the CIT

Ever since the 9/11 attacks in the US and the increased hostility of terrorist organizations, responding to and preventing security threats has become more important than ever. Detecting potential terrorists is difficult, because in many cases the critical information (e.g., location of the planned terror attack, names of the individuals involved) is not available to the investigators. In such cases, a modified version of the CIT, labeled the searching CIT (SCIT; Osugi, 2011, 2018), has been proposed. Several studies using the SCIT with groups of individuals sharing the critical items revealed that it has potential (Breska, Ben-Shakhar, & Gronau, 2012; Breska, Zaidenberg, Gronau, & Ben-Shakhar, 2014; Meijer, Bente, Ben-Shakhar, & Schumacher, 2013; Meijer, Smulders, & Merckelbach, 2010). It provides an opportunity to detect and reveal information that is not yet known to the investigators, which could be used to prevent malicious intents or find a hostage. However, future research is required to further validate the SCIT and reveal the scope of its prospective usage.

Conclusion

In short, the CIT is a method based on solid scientific principles that might resolve difficulties encountered in stress-based lie detection methods. When the appropriate items from the crime scene are selected, and intermixed with equally plausible irrelevant options, sound results can be found using memory detection. It remains important to prevent disclosure of these items during the course of the investigation, as to prevent false positives. However, as with all forensic evidence, guilt should not be solely inferred based on the CIT. Instead, memory detection can offer scientifically valid guidance on how to proceed with suspects and thereby aid in the search for the true culprit.

References

Balloun, K. D., & Holmes, D. S. (1979). Effects of repeated examinations on the ability to detect guilt with a polygraphic examination: A laboratory experiment with a real crime. *Journal of Applied Psychology, 64*(3), 316–322.

Ben-Shakhar, G. (1977). A further study on the dichotomization theory in detection of information. *Psychophysiology, 14*(4), 408–413.

Ben-Shakhar, G. (1991). Clinical judgment and decision-making in CQT-polygraphy. *Integrative Physiological and Behavioral Science, 26*(3), 232–240.

Ben-Shakhar, G. (2011). Countermeasures. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and applications of the Concealed Information Test* (pp. 200–214). Cambridge, UK: Cambridge University Press.

Ben-Shakhar, G. (2012). Current research and potential applications of the Concealed Information Test: An overview. *Frontiers in Psychology, 3,* 342.

Ben-Shakhar, G., Bar-Hillel, M., & Lieblich, I. (1986). Trial by polygraph: Scientific and juridical issues in lie detection. *Behavioral Sciences & the Law, 4*(4), 459–479.

Ben-Shakhar, G., & Dolev, K. (1996). Psychophysiological detection through the guilty knowledge technique: Effect of mental countermeasures. *Journal of Applied Psychology, 81*(3), 273–281.

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology, 88*(1), 131–151.

Ben-Shakhar, G., & Furedy, J. J. (1990). *Theories and applications in the detection of deception: A psychophysiological and international perspective.* New York, NY: Springer-Verlag.

Ben-Shakhar, G., & Nahari, T. (2018). The external validity of studies examining the detection of concealed knowledge using the Concealed Information Test. In J. P. Rosenfeld (Ed.), *Detecting concealed information and deception* (pp. 59–76). Cambridge, MA: Academic Press.

Blackstone, W. (1830). *Commentaries on the laws of England* (Vol. 2). New York, NY: Collins & Hannay.

Blakemore, S. J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: Where are we heading? *Trends in Cognitive Sciences, 8*(5), 216–222.

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology, 46*(1), 1–11.

Bradley, M. T., Barefoot, C. A., & Arsenault, A. M. (2011). Leakage of information to innocent suspects. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 187–199). Cambridge, UK: Cambridge University Press.

Bradley, M. T., & Janisse, M. P. (1981). Accuracy demonstrations, threat, and the detection of deception: Cardiovascular, electrodermal, and pupillary measures. *Psychophysiology, 18*(3), 307–315.

Bradley, M. T., & Warfield, J. F. (1984). Innocence, information, and the Guilty Knowledge Test in the detection of deception. *Psychophysiology, 21*(6), 683–689.

Breska, A., Ben-Shakhar, G., & Gronau, N. (2012). Algorithms for detecting concealed knowledge among groups when the critical information is unavailable. *Journal of Experimental Psychology: Applied, 18*(3), 292–300.

Breska, A., Zaidenberg, D., Gronau, N., & Ben-Shakhar, G. (2014). Psychophysiological detection of concealed information shared by groups: An empirical study of the searching CIT. *Journal of Experimental Psychology: Applied, 20*(2), 136–146.

Carmel, D., Dayan, E., Naveh, A., Raveh, O., & Ben-Shakhar, G. (2003). Estimating the validity of the Guilty Knowledge Test from simulated experiments: The external validity of mock crime studies. *Journal of Experimental Psychology: Applied, 9*(4), 261–269.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. Yuille (Ed.), *Credibility assessment* (pp. 51–70). Dordrecht, NL: Kluwer Academic.

Elaad, E. (1990). Detection of guilty knowledge in real-life criminal investigations. *Journal of Applied Psychology, 75*(5), 521–529.

Elaad, E., & Ben-Shakhar, G. (1991). Effects of mental countermeasures on psychophysiological detection in the Guilty Knowledge Test. *International Journal of Psychophysiology, 11*(2), 99–108.

Elaad, E., & Ben-Shakhar, G. (2009). Countering coutermeasures in the Concealed Information Test using covert respiration measures. *Applied Psychophysiology and Biofeedback, 34*(3), 197–208.

Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real-life criminal Guilty Knowledge Tests. *Journal of Applied Psychology, 77*(5), 757–767.

Gamer, M. (2011). Detecting concealed information using autonomic measures. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 27–45). Cambridge, UK: Cambridge University Press.

Gamer, M., Kosiol, D., & Vossel, G. (2010). Strength of memory encoding affects physiological responses in the guilty actions test. *Biological Psychology, 83*(2), 101–107.

Gamer, M., Verschuere, B., Crombez, G., & Vossel, G. (2008). Combining physiological measures in the detection of concealed information. *Physiology and Behavior, 95*(3), 333–340.

Geven, L. M., Ben-Shakhar, G., Kindt, M., & Verschuere, B. (2018). Memory-based deception detection: Extending the cognitive signature of lying from instructed to self-initiated cheating. *Topics in Cognitive Science.*

Geven, L. M., klein Selle, N., Ben-Shakhar, G., Kindt, M., & Verschuere, B. (2018). Self-initiated versus instructed cheating in the physiological Concealed Information Test. *Biological Psychology, 138*, 146–155.

Gustafson, L. A., & Orne, M. T. (1963). Effects of heightened motivation on the detection of deception. *Journal of Applied Psychology, 47*(6), 408–411.

Gustafson, L. A., & Orne, M. T. (1965). Effects of perceived role and role success on the detection of deception. *Journal of Applied Psychology, 49*(6), 412–417.

Honts, C. R., Devitt, M. K., Winbush, M., & Kircher, J. C. (1996). Mental and physical countermeasures reduce the accuracy of the Concealed Knowledge Test. *Psychophysiology, 33*(1), 84–92.

Horvath, F. (1978). An experimental comparison of the psychological stress evaluator and the galvanic skin response in detection of deception. *Journal of Applied Psychology, 63*(3), 338–344.

Horvath, F. (1979). Effect of different motivational instructions on detection of deception with the psychological stress evaluator and the galvanic skin response. *Journal of Applied Psychology, 64*(3), 323–330.

Iacono, W. G. (1991). Can we determine the accuracy of polygraph tests? In J. R. Jennings, P. K. Ackles, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 4, pp. 201–207). London, UK: Kingsley.

Kassin, S. M., Bogart, D., & Kerner, J. (2012). Confessions that corrupt: Evidence from the DNA exoneration case files. *Psychological Science, 23*(1), 41–45.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, E. H., & Ben-Shakhar, G. (2016). Orienting versus inhibition in the Concealed Information Test: Different cognitive processes drive different physiological measures. *Psychophysiology, 53*(4), 579–590.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, E. H., & Ben-Shakhar, G. (2017). Unraveling the roles of orienting and inhibition in the Concealed Information Test. *Psychophysiology, 54*(4), 628–639.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, E., Nahari, T., & Ben-Shakhar, G. (2017). Memory detection: The effects of emotional stimuli. *Biological Psychology, 129*, 25–35.

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry, 58*(8), 605–613.

Kugelmass, S., & Lieblich, I. (1966). Effects of realistic stress and procedural interference in experimental lie detection. *Journal of Applied Psychology, 50*(3), 211–216.

Lieblich, I., Kugelmass, S., & Ben-Shakhar, G. (1970). Efficiency of GSR detection of information as a function of stimulus set size. *Psychophysiology, 6*(5), 601–608.

Lieblich, I., Naftali, G., Shmueli, J., & Kugelmass, S. (1974). Efficiency of GSR detection of information with repeated presentation of series of stimuli in two motivational states. *Journal of Applied Psychology, 59*(1), 113–115.

Luria, A. (1932). *The nature of human conflicts.* New York, NY: Liveright.

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43*(6), 385–388.

Lykken, D. T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology, 44*(4), 258–262.

Lykken, D. T. (1974). Psychology and the lie detector industry. *American Psychologist, 29*(10), 725–739.

Lykken, D. T. (1991). Why (some) Americans believe in the lie detector while others believe in the Guilty Knowledge Test. *Integrative Physiological and Behavioral Science, 26*(3), 214–222.

Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector* (2nd ed.). New York, NY: Plenum.

Madan, C. R., & Singhal, A. (2012). Using actions to enhance memory: Effects of enactment, gestures, and exercise on human memory. *Frontiers in Psychology, 3*, 507.

Marston, W. M. (1920). Reaction-time symptoms of deception. *Journal of Experimental Psychology, 3*(1), 72–87.

Meijer, E. H., Bente, G., Ben-Shakhar, G., & Schumacher, A. (2013). Detecting concealed information from groups using a dynamic questioning approach: Simultaneous skin conductance measurement and immediate feedback. *Frontiers in Psychology, 4*, 68.

Meijer, E. H., klein Selle, N., Elber, L., & Ben-Shakhar, G. (2014). Memory detection with the Concealed Information Test: A meta-analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology, 51*(9), 879–904.

Meijer, E. H., Smulders, F. T., & Merckelbach, H. L. (2010). Extracting concealed information from groups. *Journal of Forensic Sciences, 55*(6), 1607–1609.

Mohamed, F. B., Faro, S. H., Gordon, N. J., Platek, S. M., Ahmad, H., & Williams, J. M. (2006). Brain mapping of deception and truth telling about an ecologically valid situation: Functional MR imaging and polygraph investigation—Initial experience. *Radiology, 238*(2), 679–688.

Nahari, G., & Ben-Shakhar, G. (2011). Psychophysiological and behavioral measures for detecting concealed information: The role of memory for crime details. *Psychophysiology, 48*(6), 733–744.

Nahari, T., Breska, A., Elber, L., Klein Selle, N., & Ben-Shakhar, G. (2017). The external validity of the Concealed Information Test: The effect of choosing to commit a mock crime. *Applied Cognitive Psychology, 31*(1), 81–90.

Osugi, A. (2011). Daily application of the CIT in Japan. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 253–275). Cambridge, UK: Cambridge University Press.

Osugi, A. (2018). Field findings from the Concealed Information Test in Japan. In J. P. Rosenfeld (Ed.), *Detecting concealed information and deception* (pp. 97–121). Cambridge, MA: Academic Press.

Patrick, C. J., & Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology, 76*(2), 229–238.

Pavlov, I. (1927). *Conditioned reflexes: An investigation of the psychological activity of the cerebral cortex.* New York, NY: Oxford University Press.

Pennebaker, J. W., & Chew, C. H. (1985). Behavioral—Inhibition and electrodermal activity during deception. *Journal of Personality and Social Psychology, 49*(5), 1427–1433.

Peth, J., Suchotzki, K., & Gamer, M. (2016). Influence of countermeasures on the validity of the Concealed Information Test. *Psychophysiology, 53*(9), 1429–1440.

Peth, J., Vossel, G., & Gamer, M. (2012). Emotional arousal modulates the encoding of crime-related details and corresponding physiological responses in the Concealed Information Test. *Psychophysiology, 49*(3), 381–390.

Podlesny, J. A., & Raskin, D. C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. *Psychophysiology, 15*(4), 344–359.

Reid, J. E. (1945). Simulated blood pressure responses in lie-detector tests and a method for their detection. *Journal of Criminal Law and Criminology, 36*, 201–214.

Rosenfeld, J. P., Oszan, I., & Ward, A. C. (2017). P300 amplitude at Pz and N200/N300 latency at F3 differ between participants simulating suspect versus witness roles in a mock crime. *Psychophysiology, 54*(4), 640–648.

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess "guilty knowledge". *Journal of Applied Psychology, 85*(1), 30–37.

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences, 12*(2), 48–53.

Sokolov, E. N. (1963). *Perception and the conditional reflex.* New York, NY: Macmillan.

Suchotzki, K., Verschuere, B., Peth, J., Crombez, G., & Gamer, M. (2015). Manipulating item proportion and deception reveals crucial dissociation between behavioral, autonomic and neural indices of concealed information. *Human Brain Mapping, 36*(2), 427–439.

Suchotzki, K., Verschuere, B., Van Bockstaele, B., Ben-Shakhar, G., & Crombez, G. (2017). Lying takes time: A meta-analysis on reaction time measures of deception. *Psychological Bulletin, 143*(4), 428–453.

Trovillo, P. V. (1939). A history in lie detection. *Journal of Criminal Law and Criminology, 29*(6), 848–881.

Verschuere, B., Ben-Shakhar, G., & Meijer, E. (Eds.). (2011). *Memory detection: Theory and application of the Concealed Information Test.* Cambridge, UK: Cambridge University Press.

Verschuere, B., Crombez, G., Koster, E. H., Van Bockstaele, B., & De Clercq, A. (2007). Startling secrets: Startle eye blink modulation by concealed crime information. *Biological Psychology, 76*(1–2), 52–60.

Verschuere, B., Meijer, E., & De Clercq, A. (2011). Concealed information under stress: A test of the orienting theory in real-life police interrogations. *Legal and Criminological Psychology, 16*(2), 348–356.

Vrij, A. (2004). Guidelines to catch a liar. In P. Granhag & L. Stromwall (Eds.), *The detection of deception in forensic contexts* (pp. 287–314). Cambridge, UK: University Press.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester, UK: Wiley.

Zaitsu, W. (2016). External validity of Concealed Information Test experiment: Comparison of respiration, skin conductance, and heart rate between experimental and field card tests. *Psychophysiology, 53*(7), 1100–1107.

True and False Intentions:
A Science of Lies About the Future

author block

Eric Mac Giolla and Pär Anders Granhag

Research on deception detection has traditionally focused on statements about the past. The focus is on answers to questions such as "where *were* you yesterday?" and "what *did* you do there?" In the current chapter, we provide an overview of the burgeoning field of true and false intentions. Here, the focus changes to statements about the future. Specifically, research concerns statements about one's own future behavior, that is, statements of intent. Put differently, the focus changes to answers of questions such as "where *will* you go tomorrow?" and "what *will* you do there?" This subtle change of tense may seem trivial. However, as will be shown, the topic of true and false intentions brings with it difficulties and opportunities distinct from traditional research on truths and lies about past events. In the current chapter, we provide an overview of the topic of true and false intentions, summarize the extant research, and describe our recent attempts at creating a framework to parsimoniously account for the emerging findings of this novel field of deception detection.

E. Mac Giolla (✉) · P. A. Granhag
University of Gothenburg, Gothenburg, Sweden
e-mail: Erik.Mac.Giolla@psy.gu.se

P. A. Granhag
Norwegian Police University College, Oslo, Norway
e-mail: pag@psy.gu.se

Definitions and First Studies

When defining true and false intentions, deception researchers have availed of more basic social psychological findings (Granhag, 2010). According to such research, an intention is composed of three necessary parts: a *goal*, that has come with some degree of *reasoning*, that one is *committed* to carry out (Malle & Knobe, 1997, 2001). It is this commitment that distinguishes an intention from related concepts such as desires. That is, although we can desire any number of things, such desires only become intentions once we decide to carry out the actions necessary to fulfill them (Malle & Knobe, 2001). Furthermore, when we discuss intentions from the perspective of deception detection, we are concerned with stated intentions. Based on Malle and Knobe's definition, a statement of true intent refers to a future action or goal which a speaker is committed to carry out. In contrast, a statement of false intent comes with no such commitment. Instead, a statement of false intent refers to a future action or goal which a speaker claims, but does not in fact intend to carry out. A common reason for a false intention is as a cover story, used to mask one's genuine intention. Such a cover story was told by 22-year-old Matti Saari in the western part of Finland, September 22, 2008 ("Finnish college gunman kills 10", 2008). Matti Saari was asked by the police about his new weapons license and the video-clips he recently had posted on the Internet, showing him firing his Walther P22 Target. He assured the police that it was for recreation and leisure. But this was just a cover story. The next morning Saari dressed in black, went to his school, and killed nine students and one teacher. Then, he took his own life.

The content of intentions can vary dramatically. They can refer to concrete tasks in the near future (e.g., to clean the living room this evening) or to more abstract tasks in the distant future (e.g., to live healthier). For practical research reasons, the majority of studies on the topic have focused on relatively simple intentions of tasks to be performed in the near future. For example, the first study on true and false intentions focused on statements of intent concerning future trips (Vrij, Granhag, Mann, & Leal, 2011). Travelers in an airport were approached to participate in the study. Half of the participants provided a truthful description of their travel plans and hence provided true intentions. The other half lied about their travel plans and hence provided false intentions. Specifically, they provided false information about the destination and activities of their trip. The true and false statements of intent did not differ in terms of length and detail. However, the true intentions were rated as more plausible. Furthermore, when independent judgers were asked to assess if the statements were true or false, they made correct decisions approximately 70% of the time. Almost identical results were found in a follow-up study where military personnel provided true or false statements of intent about a mock future mission (Vrij, Leal, Mann, & Granhag, 2011). Statements of true intent were again deemed as more plausible,

and independent judgers correctly assessed statements as true or false around 70% of the time.

The relatively high accuracy rates of these two studies stand in contrast to the modest rates typically observed in deception studies (Bond & DePaulo, 2006). These two studies, however, are largely explorative and offer little theory or insight into the topic of true and false intentions. Later studies have approached the topic in a more systematic manner, albeit from many different starting points. In a recent summary of the field, we broadly categorize these different approaches into (1) studies that extend traditional approaches of deception detection to the field of true and false intentions and (2) studies that develop new approaches, which are grounded in theory and research specific to intention situations (Mac Giolla, Granhag, & Vrij, 2015). These approaches will be described in turn (for a more detailed review of the field, see Granhag & Mac Giolla, 2014).

EXTENDING TRADITIONAL APPROACHES TO INTENTION SITUATIONS

Physiological Studies

Deception detection techniques based on physiological measures have a long history in research and practice (Trovillo, 1939). A handful of deception detection techniques based on physiological measures have been extended to situations of intent with varying success. These techniques vary in terms of what is measured as well as their theoretical underpinnings. They include studies based on eye-tracking and thermal imaging, as well as studies on the concealed information test (CIT) that measure skin conductance and neural activity. Here, we will focus on studies on thermal imaging and CIT (for research on intentions and eye-tracking, see Mann, Vrij, Leal, et al., 2012; Mann, Vrij, Nasholm, et al., 2012).

Proponents of thermal imaging approaches argue that lying results in an increased physiological arousal which can be uncovered by measuring body temperature. Support for this position was found in a study about lying about past events (Pavlidis, Eberhardt, & Levine, 2002). Generalizing from their results, the authors suggested that the method may also be suitable to distinguish true from false statements of intent. However, later research testing this proposal suggests otherwise (Warmelink et al., 2011). The exact reasons why the thermal imaging technique did not generalize to an intent situation are unknown. It could potentially highlight a boundary condition of the approach. Alternatively, it could highlight a more general problem of deception detection techniques based on the idea that lying increases arousal (for more general criticisms of arousal-based approaches to deception, see DePaulo et al., 2003).

Studies examining concealed information tests (CITs) have performed considerably better when extended to intention contexts. CITs aim to

distinguish between the absence and the presence of information in some-one's long-term memory (Verschuere, Ben-Shakhar, & Meijer, 2011). Theoretically, the CIT is based on the well-established orienting response—a specific reflexive reaction produced to stimuli that has significance to an indi-vidual (e.g., one's own name) (Meixner & Rosenfeld, 2011; Sokolov, 1963). The CIT has shown that similar responses occur when people are confronted with information about knowledge they possess, but are trying to conceal. In an archetypal example, a murder suspect is presented with pictures of sev-eral potential murder weapons. The innocent suspect should show the same reaction to all pictures. In contrast, the guilty suspect should show an orient-ing response for the picture of the actual murder weapon. This response can in turn be accurately gauged by measuring skin conductance (Verschuere & Meijer, 2014) or neural activity in the form of event-related potentials (ERPs) (Rosenfeld et al., 1988).

CITs based on both skin conductance and ERPs have been successfully extended to situations of intent (Meijer, Verschuere, & Merckelbach, 2010; Meixner & Rosenfeld, 2011). In the study by Meijer et al., for example, par-ticipants planned a mock crime or a non-criminal task. Half of those who planned the mock crime actually carried it out, while the other half were stopped before they were able to perform the task. All participants performed a CIT related to the mock crime. Questions during the CIT referred to such aspects as the object that participants stole or were planning to steal. Innocent suspects who were not informed about the details of the crime did not show a heightened response for the crime-related target answers. In contrast, both those who actually carried out the task and those who merely planned to carry out the task showed the expected orienting response, to tar-get answers.

Reaction Time Studies

A number of deception detection techniques based on reaction time measures have been developed. Two of these, the autobiographical Implicit Association Test (aIAT) (Sartori, Agosta, Zogmaister, Ferrara, & Castiello, 2008) and the Sheffield Lie Test (Spence et al., 2001), have been extended to intention sit-uations with promising results. Here, we will focus on the aIAT because it is, thus far, the only deception detection technique that attempts to distinguish not only true intentions from false ones, but also true intentions from related concepts such as desires (for research on the Sheffield Lie Test and intentions, see Suchotzki, Verschuere, Crombez, & De Houwer, 2013).

In the standard aIAT, autobiographical statements appear one by one on a computer screen. These statements come in four forms: verifiably true statements (e.g., I am in front of a computer); verifiably false statements (e.g., I am on a beach); innocent statements related to the crime in question (e.g., I am innocent of the crime); and guilty statements related to the crime

in question (e.g., I am guilty of the crime). Participants are then required to classify these statements based on designated keyboard keys as quickly as possible. Faster response times are seen for truth tellers (innocent suspects), compared to liars (guilty suspects) in trials when true statements and innocent statements are designated the same key. In contrast, faster response times are seen for liars, compared to truth tellers, in trials when true statements and guilty statements are designated the same key (Sartori et al., 2008).

The same basic idea has been used in studies where the aIAT has been applied to intention situations. Again, participants categorize autobiographical statements that appear on a screen. As with the standard aIAT, verifiably true or false statements are classified by pressing separate keys. The difference is that in intention studies verifiably true and false statements are coupled with true and false statements of intent, rather than statements pertaining to past guilt or innocence (Agosta, Castiello, Rigoni, Lionetti, & Sartori, 2011). For example, in the study by Agosta et al., the statements of intent concerned, among other things, where the participant intended to sleep that night. During congruent trials, statements of true intent were classified by pressing the key designated to verifiably true statements, and statements of false intent were classified by pressing the key designated to verifiably false statements. During incongruent trials, statements of true intent were classified by pressing the key designated to verifiably false statements, while statements of false intent were classified by pressing the key designated to verifiably true statements. As predicted, faster reaction times (RTs) were seen during congruent versus incongruent trials.

In a follow-up experiment, Agosta et al. (2011) compared statements of hope to statements of intent. A statement of hope was the desired outcome that was not likely to be carried out or to occur (e.g., I hope to win the lottery). Such statements are therefore similar to "desires", as described by Malle and Knobe (1997), which, unlike intentions, do not come with a commitment to be carried out. In brief, results showed that statements of hope showed slower RTs when matched with true statements, compared to the RTs when statements of intent were matched to true statements. In other words, statements of intent were more closely represented as true, likely because the participants believed that these would eventually lead to action. Due to the elusive nature of intentions, such findings are crucial for validating the broader project of true and false intentions. The boundaries of this field of inquiry have been pushed even further in a recent study (Zangrossi, Agosta, Cervesato, Tessarotto, & Sartori, 2015). In brief, Zangrossi and colleagues examined the possibilities of an amended version of the aIAT to distinguish between true and false statements about past intentions. In other words, they examined how to distinguish between one's intentions to perform an act that has already been committed. The results showed that aIAT could detect true past intentions, and as such, is one of the few lie detection studies to address the highly practical issue of determining *mens rea*.

Strategic Interviewing

Strategic interviewing methods represent one of the most promising developments in deception detection in recent years (Vrij & Granhag, 2012). Despite a variety of different approaches, strategic interviewing techniques rest on the same basic idea. Simply put, due to the paucity of reliable deception cues, the interviewer (or lie-catcher) must actively attempt to elicit cues through the use of strategic questioning (Vrij, 2015). At least two of these strategic interviewing techniques have been extended to intention situations, with varying degrees of success.

First, the strategic use of evidence (SUE) technique (Hartwig, Granhag, Strömwall, & Kronkvist, 2006) has been extended to intention situations (Clemens, Granhag, & Strömwall, 2011). The SUE technique builds on the different counter-interrogation strategies of truth tellers and liars. In brief, truth tellers are much more forthcoming with information, particularly with regard to potentially incriminating details. Accordingly, by asking open questions, exhausting alternative explanations, and by disclosing the available evidence in a tactical manner, these differing strategies should result in truth tellers providing statements that correspond more with the available evidence compared to liars. Put differently, liars should be more likely to avoid and/or deny critical information, thereby increasing the chances of providing statement evidence inconsistencies (Granhag & Hartwig, 2015). The exact same reasoning has been applied with great success to situations of intent. In the study by Clemens et al. (2011), those with a true intention had planned to perform an innocuous task, while those with a false intention had planned to perform a mock-crime. The interviewer had evidence that related to the planning that those with a true and false intention had conducted in preparation for their future tasks (e.g., web browser history). When interviewed about this planning phase, those who were to perform a mock crime (i.e., those stating false intentions) withheld more information, resulting in more statement evidence inconsistencies, compared to those with a true intention.

Second, the reverse order technique (Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2007) has been examined in situations of intent (Fenn, McGuire, Langben, & Blandón-Gitlin, 2015). The reverse order technique is one tactic of Colwell et al.'s strategic interviewing method, Assessment Criteria Indicative of Deception (ACID). In brief, it builds on the idea that truth tellers will have a memory of the event they describe, while liars will rely on prepared lie scripts. In contrast to the multi-encoded and flexible memory of a real event, lie scripts are primarily verbal and rigid in nature. This puts liars in a difficult position when they are required to provide their answers in reverse order. Accordingly, liars have been shown to provide shorter and less detailed answers compared to truth tellers for such questions. However, unlike the SUE technique, reverse order questioning has shown to be less effective when used to distinguish between statements of true and false intent (Fenn et al., 2015). Specifically, no differences in detail and length were

found between descriptions of true and false intentions provided in reverse order. The authors' explanation was that describing even true intentions is a difficult task. Hence, the added difficulty associated with providing answers in reverse order may be similarly demanding for truth tellers and liars. Such results highlight the difficulty in simply translating techniques developed to distinguish true and false statements about past events to intention contexts. This provides the impetus for the topic we turn to next: novel, intention-specific, deception detection approaches.

Intention-Specific Deception Detection Approaches

Research on intention-specific approaches begins with the observation that intentions are closely related to a host of other psychological constructs (Malle, Moses, & Baldwin, 2001). More specifically, the focus has been on the constructs that typically accompany the formation of (true) intentions. Research to date has focused on constructs such as planning (Sooniste, Granhag, Strömwall, & Vrij, 2016), episodic future thought (EFT) (Knieps, Granhag, & Vrij, 2013a), goals (Ask, Granhag, Juhlin, & Vrij, 2013), and spontaneous thoughts (Mac Giolla, Granhag, & Ask, 2017b). Due to the relatively large number of studies so far conducted on planning and EFT, we will limit our discussion to these studies (for a more comprehensive overview, see Granhag & Mac Giolla, 2014; Mac Giolla et al., 2015).

Intentions and Planning

Intentions and planning are closely related concepts. This is already clear in the definition of intentions provided by Malle and Knobe (1997, 2001), who explain that intentions are typically preceded by some degree of reasoning, which can be seen as rudimentary plans. Intentions, however, can also be a catalyst for more detailed plans. This is because plans help us carry out our intentions (Mumford, Schultz, & Van Doorn, 2001). The clearest example of this comes from research on implementation intentions (Gollwitzer & Brandstätter, 1997). In this research, a distinction is made between goal intentions and implementation intentions. Goal intentions refer to the "what" (e.g., I intend to lose weight). Implementation intentions refer to concrete plans that outline the steps needed to carry out the goal intention. Put differently, as well as the "what", implementation intentions also include the "where", "when", and "how" (e.g., I intend to lose weight by jogging, in the park beside my house, weekday mornings before work). Research indicates that these simple, but concrete plans, drastically improve people's ability to fulfill their intentions (Gollwitzer & Sheeran, 2006).

How does this relate to true and false intentions? In short, research indicates that truth tellers are, in comparison to liars, more likely to engage in the in-depth planning required to attain their intentions. Sooniste (2015)

explains this idea in terms of bounded rationality (Simon, 1978). That is, truth tellers, since their stated intentions are genuine, are motivated to carry them out. This motivation should lead to more optimizing behavior, such as in-depth planning. In contrast, liars do not actually intend to carry out their intentions. Hence, they should be less motivated to engage in the resource draining behavior of planning. Rather, liars will satisfice. They will prepare statements of false intent or cover stories that are sufficiently detailed to be believed. Liars are of course motivated to be believed, and hence, they should produce detailed cover stories. However, it is perhaps less likely that they will go to the next step and engage in the concrete planning activities that would be required to actually achieve their stated intentions. A growing body of research suggests that this may be the case.

For example, in one study, when asked to describe the planning that went into their intentions, truth tellers provided longer and more detailed descriptions of the planning phase compared to liars (Sooniste, Granhag, Knieps, & Vrij, 2013). Similar results were found by Mac Giolla and Granhag (2015), who found that although truth tellers provided longer descriptions for both their intentions and the planning phase compared to liars, these differences were considerably larger for the descriptions of the planning phase. Related work also indicates that truth tellers produce better plans than liars. For instance, truth tellers tend to describe more efficient, concrete, and flexible plans than liars (Mac Giolla, Granhag, & Liu-Jönsson, 2013), all of which are considered to be indicative of good planning behavior (e.g., Mumford et al., 2001). This idea of concrete plans also fits well with the results of another study (Warmelink, Vrij, Mann, Jundi, & Granhag, 2012). In the study by Warmelink et al., participants were interviewed about a future trip. Half of the participants were actually to go on this trip and thus provided a true intention. The other half only claimed they were to go on the trip and thus provided a false intention. Participants were asked a number of specific and general questions about the trip and the planning that went into the trip. Truth tellers provided more detailed statements, but only for questions concerning concrete aspects of planning, such as which transportation they were to use.

Episodic Future Thoughts

If planning can be seen as a typical concomitant of the formation of true intentions, then episodic future thoughts (EFTs) can be seen as a typical concomitant of planning. EFTs refer to the mental images one experiences when thinking about the future (Szpunar, 2010). Specifically, they refer to the mental representation of specific autobiographical future events. They are often accompanied by vivid imagery and can be seen as an analogue of autobiographical memory. The parallels between EFTs and autobiographical memories are further highlighted by research showing that mental simulations

of future events are largely constructed from episodic memories (Schacter, Addis, & Buckner, 2008).

With regard to the topic of true and false intentions, research has shown that truth tellers experience EFTs related to their stated intentions more often and experience them as more vivid compared to liars (Granhag & Knieps, 2011). This finding has been replicated in a series of studies by Knieps and colleagues (Knieps et al., 2013a; Knieps, Granhag, & Vrij, 2013b, 2014). Although these differences refer to the subjective experiences of a mental image, they can nonetheless aid in the task of deception detection. Specifically, such research can provide a base from which to build strategic interviewing methods.

One of the more straightforward strategic interviewing methods is the unanticipated questions approach (Vrij et al., 2009). The basic idea is to ask questions which a truth teller should be capable of answering with relative ease, but for which a liar has not prepared a response. Unprepared lies are easier to detect than prepared lies (DePaulo et al., 2003). Hence, by asking unanticipated questions, it should be possible to elicit differences between truth tellers and liars. When applied to past events, these unanticipated questions typically concern information that truth tellers should have a clear memory of, but which liars may be less likely to have considered in their lie scripts (e.g., spatial information, Vrij et al., 2009). Alternatively, unanticipated questions can avail of unanticipated response formats. These should enhance a true memory of an event, but hamper liars' recall of a rigid lie-script. An example of unanticipated questions of this form includes the reverse order questioning mentioned above (Colwell et al., 2007).

The problem when applying these forms of unanticipated questions to intention situations is that they rely on the truth teller having a memory of an event. With intentions, of course, there is no memory, since the event has yet to take place. This can account for the problems encountered by Fenn et al. (2015), noted above, when trying to extend the reverse order technique to situations of intent. However, by focusing on tendencies and consequences associated with the formation of a true intention, such as planning or EFT, we argue it is still possible to develop themes which are suitable for unanticipated questions. For example, in the series of studies by Knieps and colleagues (Knieps et al., 2013a, 2013b, 2014), differences were observed between truth tellers and liars, not only in their experience of EFTs, but also how they answered questions about these images during an interview. Importantly, in this interview liars were trying to convince the interviewer that their stated intentions were true. Results showed that truth tellers were significantly more likely than liars to admit to having experienced EFTs concerning their stated intentions, and, when asked to describe them, typically provided longer descriptions of their EFTs than liars. In a follow-up study, participants were also asked to draw these mental images (Calderon, Mac Giolla, Ask, & Granhag, 2018). In the study by Calderon et al. drawings of

EFTs of true intentions were rated as significantly more concrete than drawings of EFTs of false intentions. Another study demonstrated how more specific differences in content can be elicited by probing these mental images with more precise questions (Warmelink, Vrij, Mann, & Granhag, 2013). Specifically, Warmelink et al.'s findings showed that truth tellers provided significantly more spatial and temporal details of their mental images compared to liars. Again, the explanation for such findings is that truth tellers have a clearer and more vivid future image than liars and are therefore in a better position to provide such specific details.

To reiterate, we argue that the typical tendencies associated with the formation of a true intention can be used as a base for developing strategic interviewing methods. As will be outlined in the next section, we believe that EFTs and planning represent just the tip of the iceberg of such typical tendencies.

A Suggested Framework for True and False Intentions

In a recent article, we suggest a framework to integrate the growing body of work on intention-specific approaches to deception (Mac Giolla, Granhag, & Ask, 2017a; see also, Mac Giolla, 2016). Our aim was to find a common factor that could both account for past research and develop new approaches. We suggest that "active goals" are central for this task. According to the definition of intentions provided above, true intentions consist of a goal that comes with a commitment to act. Put differently, true intentions imply the activation of a goal. Active goals have, in turn, a plethora of predictable consequences on behavior—a finding reiterated throughout the history of psychology (Ajzen, 1991; Ajzen & Fishbein, 1969; Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001; Förster, Liberman, & Friedman, 2007; Lewin, 1935; Martin & Tesser, 2009; Tolman, 1932). Among other things, active goals affect our judgments (Strachman & Gable, 2006), memory (Zeigarnik, 1939), attention (Moskowitz, 2002), and evaluations (Ferguson & Bargh, 2004). Many of these effects are automatic (Bargh et al., 2001). But active goals also have more deliberate consequences, such as reasoned planning (Ajzen, 2012). These consequences of active goals are not haphazard. Rather, they represent functional behaviors that aid in goal attainment. Simply put, these consequences help us achieve our goals (for a more detailed discussion, see Mac Giolla et al., 2017a). Crucially, those stating a false intention are not committed to carry out the goal of their stated intention. In other words, false intentions do not lead to active goals. For this reason, false intentions should be less likely to be accompanied by the typical consequences of active goals associated with the formation of true intentions.

As an example, consider the study by Ask et al. (2013). The study is based on the finding that objects are automatically and implicitly evaluated based on their utility for an active goal. Specifically, objects that aid in goal attainment are evaluated positively, while objects that hinder goal attainment are

evaluated negatively (De Houwer, 2009; Ferguson, 2007; Ferguson & Bargh, 2004). In the study by Ask et al., truth tellers planned an innocuous task—a shopping trip in a nearby shopping mall. Liars planned a mock crime, but were also told to prepare a cover story. The cover story concerned the same task that truth tellers planned for. Hence, the shopping trip was the true intention for truth tellers and the false intention for liars. After planning their tasks, but before they could carry them out, participants performed an evaluative priming task—an established implicit measure of evaluations (Fazio, Jackson, Dunton, & Williams, 1995; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). The words evaluated were facilitative of the shopping task (e.g., "cash register"). Truth tellers showed a positive implicit evaluation of the target words, in accordance with what is expected of having an active shopping goal. In contrast, liars showed a neutral evaluation of the target words, in accordance with what would be expected when one does not have an active shopping goal.

From a functional perspective, the findings of Ask et al. (2013) make perfect sense. Those with a true intention have an active shopping goal. Evaluating stimuli based on this goal lets them know what to approach or avoid in order to carry out their intentions and achieve their goals. In contrast, those stating a false intention do not have an active shopping goal. Therefore, evaluating objects based on this non-active goal makes little sense. Put differently, since they do not plan to carry out the shopping task, there is no functional value in knowing what to approach and avoid. A similar functional interpretation can also account for the findings on planning and EFTs described above. Both planning and EFTs are functional consequences of active goals and hence true intentions. That is, they both aid in goal attainment. It is obvious that planning aids in goal attainment (Mumford et al., 2001), and the more concrete the plan the better (Gollwitzer, 1999). But, EFTs also aid in goal attainment. By imagining a future scenario, it is possible to mentally rehearse or practice the future event, and thereby increase chances of goal attainment (Schacter et al., 2008). These functional consequences, however, are lost on those stating a false intention. Since they are not committed to carry out their future goals, typical consequences of intentions that aid in goal attainment become redundant. Furthermore, behaviors such as planning and episodic future thinking are cognitively demanding. Considering that lying is in itself a difficult task, it makes sense that those stating a false intention will engage in such behaviors to a lower degree than truth tellers. This is exactly what the research, described above, shows: Those with a true intention produce better plans and more detailed plans, and engage in EFT to a greater extent than those with a false intention.

A focus on the functional consequences of active goals can also shed light on new avenues of research. As noted above, active goals influence, among other things, what we remember, what we attend to, and even how we judge and interpret stimuli. These consequences of active goals have, as of yet,

not been studied within the topic of true and false intentions. However, we would predict that such consequences should be stronger for those stating a true intention—since only they have an active goal. In many situations, those stating a false intention will have an ulterior (active) goal that they are trying to hide. The typical consequences of active goals should of course accompany these ulterior goals. We posit, however, that they should be less likely to accompany the non-active goals of their stated false intention. The research conducted so far shows that these expected differences between those stating a true and false intention can be uncovered by at least two means. First, the study by Ask et al. (2013) shows the potential in using implicit measures, such as evaluative priming tasks. Second, the research on planning and EFTs shows how these expected tendencies can provide the base from which to develop strategic interviewing methods viable in intention contexts.

Conclusion

Research on true and false intentions is a new area of inquiry within the field of legal psychology, and the body of research is steadily approaching a critical mass. In this chapter, we have summarized the extant research. We first reviewed research that can be seen as extensions of traditional deception detection research (physiological studies, reaction time studies, and strategic interviewing studies). We then turned to intention-specific initiatives and reviewed work on (a) intentions and planning and (b) EFTs. Finally, we introduced a theoretical framework accommodating past findings and providing signposts for future work in this emerging field. Research examining both traditional and novel approaches will likely be necessary to unravel the difficult task of distinguishing between true and false intent. The potential rewards, however, are considerable.

References

Agosta, S., Castiello, U., Rigoni, D., Lionetti, S., & Sartori, G. (2011). The detection and the neural correlates of behavioral (prior) intentions. *Journal of Cognitive Neuroscience, 23*(12), 3888–3902. https://doi.org/10.1162/jocn_a_00039.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Ajzen, I. (2012). The theory of planned behavior. In P. A. M. Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 438–459). London: Sage.

Ajzen, I., & Fishbein, M. (1969). The prediction of behavioral intentions in a choice situation. *Journal of Experimental Social Psychology, 5*(4), 400–416.

Ask, K., Granhag, P. A., Juhlin, F., & Vrij, A. (2013). Intending or pretending? Automatic evaluations of goal cues discriminate true and false intentions. *Applied Cognitive Psychology, 27*(2), 173–177. https://doi.org/10.1002/acp.2893.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology, 81*(6), 1014–1027. https://doi.org/10.1037/0022-3514.81.6.1014.

Bond, C. F., Jr. & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Calderon, S., Mac Giolla, E., Granhag, P. A., & Ask, K. (2018). Drawing what lies ahead: False intentions are more abstractly depicted than true intentions. *Applied Cognitive Psychology, 32*, 518–522. https://doi.org/10.1002/acp.3422.

Clemens, F., Granhag, P. A., & Strömwall, L. A. (2011). Eliciting cues to false intent. *Law and Human Behavior, 35*(6), 512–522. https://doi.org/10.1007/s10979-010-9258-9.

Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007). Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling, 4*(3), 167–180.

De Houwer, J. (2009). How do people evaluate objects? A brief review. *Social and Personality Psychology Compass, 3*(1), 36–48.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*(2), 229.

Fenn, E., McGuire, M., Langben, S., & Blandón-Gitlin, I. (2015). A reverse order interview does not aid deception detection regarding intentions. *Frontiers in Psychology, 6,* 1298. https://doi.org/10.3389/fpsyg.2015.01298.

Ferguson, M. J. (2007). On the automatic evaluation of end-states. *Journal of Personality and Social Psychology, 92*(4), 596. https://doi.org/10.1037/0022-3514.92.4.596.

Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: The effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology, 87*(5), 557–572. https://doi.org/10.1037/0022-3514.87.5.557.

Finnish college gunman kills 10. (2008). *BBC News.* Retrieved from http://news.bbc.co.uk.

Förster, J., Liberman, N., & Friedman, R. S. (2007). Seven principles of goal activation: A systematic approach to distinguishing goal priming from priming of non-goal constructs. *Personality and Social Psychology Review, 11*(3), 211–233. https://doi.org/10.1177/1088868307303029.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*(7), 493–503. https://doi.org/10.1037/0003-066X.54.7.493.

Gollwitzer, P. M., & Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology, 73*(1), 186.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achieve-ment: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology, 38,* 69–119. https://doi.org/10.1016/S0065-2601(06)38002-1.

Granhag, P. A. (2010). On the psycho-legal study of true and false intentions: Dangerous waters and some stepping stones. *The Open Criminology Journal, 3,* 37–43. https://doi.org/10.2174/1874917801003010037.

Granhag, P. A., & Hartwig, M. (2015). The Strategic Use of Evidence (SUE) tech-nique: A conceptual overview. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: New challenges and cognitive approaches* (pp. 231–251). Chichester, UK: Wiley.

Granhag, P. A., & Knieps, M. (2011). Episodic future thought: Illuminating the trademarks of forming true and false intentions. *Applied Cognitive Psychology, 25*(2), 274–280. https://doi.org/10.1002/acp.1674.

Granhag, P. A., & Mac Giolla, E. (2014). Preventing future crimes: Identifying mark-ers of true and false intent. *European Psychologist, 19*(3), 195–206. https://doi.org/10.1027/1016-9040/a000202.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior, 30*(5), 603–619. https://doi.org/10.1007/s10979-006-9053-9.

Knieps, M., Granhag, P. A., & Vrij, A. (2013a). Back to the future: Asking about mental images to discriminate between true and false intentions. *The Journal of Psychology, 147*(6), 619–640. https://doi.org/10.1080/00223980.2012.728542.

Knieps, M., Granhag, P. A., & Vrij, A. (2013b). Repeated visits to the future: Asking about mental images to discriminate between true and false intentions. *International Journal of Advances in Psychology, 2*(2), 93–102.

Knieps, M., Granhag, P. A., & Vrij, A. (2014). Been there before? Examining "familiarity" as a moderator for discriminating between true and false intentions. [Original Research]. *Frontiers in Psychology, 5,* 1–9. https://doi.org/10.3389/fpsyg.2014.00677.

Lewin, K. (1935). *A dynamic theory of personality*. New York: McGaw-Hill.

Mac Giolla, E. (2016). *Towards a theory of true and false intentions*. (Doctoral thesis), Departement of Psychology, University of Gothenburg, Sweden. Retrieved from https://gupea.ub.gu.se/bitstream/2077/43237/1/gupea_2077_43237_1.pdf.

Mac Giolla, E., & Granhag, P. A. (2015). Detecting false intent amongst small cells of suspects: Single versus repeated interviews. *Journal of Investigative Psychology and Offender Profiling, 12,* 142–157. https://doi.org/10.1002/jip.1419.

Mac Giolla, E., Granhag, P. A., & Ask, K. (2017a). A goal-activation framework of true and false intentions. *Applied Cognitive Psychology, 31,* 678–684. https://doi.org/10.1002/acp.3366.

Mac Giolla, E., Granhag, P. A., & Ask, K. (2017b). Task-related spontaneous thought: A novel direction in the study of true and false intentions. *Journal of Applied Research in Memory and Cognition, 6*(1), 93–103. https://doi.org/10.1016/j.jarmac.2016.04.010.

Mac Giolla, E., Granhag, P. A., & Liu-Jönsson, M. (2013). Markers of good planning behavior as a cue for separating true and false intent. *PsyCh Journal, 2*(3), 183–189. https://doi.org/10.1002/pchj.36.

Mac Giolla, E., Granhag, P. A., & Vrij, A. (2015). Discriminating between true and false intentions. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception*

detection: Current challenges and new approaches (pp. 155–173). Chichester, UK: Wiley.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33*(2), 101–121. https://doi.org/10.1006/jesp.1996.1314.

Malle, B. F., & Knobe, J. (2001). The distinction between desire and intention: A folk-conceptual analysis. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 45–68). Cambridge, MA: MIT Press.

Malle, B. F., Moses, L. J., & Baldwin, D. A. (2001). *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press.

Mann, S., Vrij, A., Leal, S., Granhag, P. A., Warmelink, L., & Forrester, D. (2012). Windows to the soul? Deliberate eye contact as a cue to deceit. *Journal of Nonverbal Behavior*, 1–11. https://doi.org/10.1007/s10919-012-0132-y.

Mann, S., Vrij, A., Nasholm, E., Warmelink, L., Leal, S., & Forrester, D. (2012). The direction of deception: Neuro-linguistic programming as a lie detection tool. *Journal of Police and Criminal Psychology, 27*(2), 160–166. https://doi.org/10.1007/s11896-011-9097-8.

Martin, L. L., & Tesser, A. (2009). Five markers of motivated behavior. In G. B. Moskowitz & H. Grant (Eds.), *Psychology of goals*. London: The Guilford Press.

Meijer, E. H., Verschuere, B., & Merckelbach, H. (2010). Detecting criminal intent with the concealed information test. *Open Criminology Journal, 3,* 44–47. https://doi.org/10.2174/1874917801003020044.

Meixner, J. B., & Rosenfeld, J. P. (2011). A mock terrorism application of the P300-based concealed information test. *Psychophysiology, 48*(2), 149–154. https://doi.org/10.1111/j.1469-8986.2010.01050.x.

Moskowitz, G. B. (2002). Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology, 38*(4), 397–404.

Mumford, M. D., Schultz, R. A., & Van Doorn, J. R. (2001). Performance in planning: Processes, requirements, and errors. *Review of General Psychology, 5*(3), 213–240. https://doi.org/10.1037/1089-2680.5.3.213.

Pavlidis, I., Eberhardt, N. L., & Levine, J. A. (2002). Human behaviour: Seeing through the face of deception. *Nature, 415*(6867), 35. https://doi.org/10.1038/415035a.

Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience, 42*(1–2), 157–161.

Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S. D., & Castiello, U. (2008). How to accurately detect autobiographical events. *Psychological Science, 19*(8), 772–780.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events. *Annals of the New York Academy of Sciences, 1124*(1), 39–60. https://doi.org/10.1196/annals.1440.001.

Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review, 68,* 1–16.

Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology, 25*(1), 545–580.

Sooniste, T. (2015). *Discriminating between true and false intentions: Questions to pose and cues to use*. (Doctoral thesis), Departement of Psychology, University of Gothenburg, Sweden. Retrieved from https://gupea.ub.gu.se/handle/2077/38724.

Sooniste, T., Granhag, P. A., Knieps, M., & Vrij, A. (2013). True and false intentions: Asking about the past to detect lies about the future. *Psychology, Crime & Law, 19*(8), 673–685. https://doi.org/10.1080/1068316x.2013.793333.

Sooniste, T., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2016). Discriminating between true and false intent among cells of suspects. *Legal and Criminological Psychology, 21*(2), 344–357. https://doi.org/10.1111/lcrp.12063.

Spence, S. A., Farrow, T. F., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. (2001). Behavioural and functional anatomical correlates of deception in humans. *NeuroReport, 12*(13), 2849–2853.

Strachman, A., & Gable, S. L. (2006). What you want (and do not want) affects what you see (and do not see): Avoidance social goals and social events. *Personality and Social Psychology Bulletin, 32*(11), 1446–1458.

Suchotzki, K., Verschuere, B., Crombez, G., & De Houwer, J. (2013). Reaction time measures in deception research: Comparing the effects of irrelevant and relevant stimulus–response compatibility. *Acta Psychologica, 144*(2), 224–231. https://doi.org/10.1016/j.actpsy.2013.06.014.

Szpunar, K. K. (2010). Episodic future thought an emerging concept. *Perspectives on Psychological Science, 5*(2), 142–162. https://doi.org/10.1177/1745691610362350.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. California: University of California Press.

Trovillo, P. V. (1939). A history of lie detection. *Journal of Criminal Law and Criminology (1931–1951), 29*, 848–881.

Warmelink, L., Vrij, A., Mann, S., & Granhag, P. A. (2013). Spatial and temporal details in intentions: A cue to detecting deception. *Applied Cognitive Psychology, 27*(1), 101–106. https://doi.org/10.1002/acp.2878.

Warmelink, L., Vrij, A., Mann, S., Jundi, S., & Granhag, P. A. (2012). The effect of question expectedness and experience on lying about intentions. *Acta Psychologica, 141*(2), 178–183. https://doi.org/10.1016/j.actpsy.2012.07.011.

Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D., & Fisher, R. P. (2011). Thermal imaging as a lie detection tool at airports. *Law and Human Behavior, 35*(1), 40–48. https://doi.org/10.1007/s10979-010-9251-3.

Verschuere, B., Ben-Shakhar, G., & Meijer, E. (2011). *Memory detection: Theory and application of the concealed information test*. Cambridge, UK: Cambridge University Press.

Verschuere, B., & Meijer, E. H. (2014). What's on your mind? Detecting concealed information. *European Psychologist, 19*, 162–171.

Vrij, A. (2015). A cognitive approach to lie detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches* (pp. 205–229). Chichester: Wiley.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*(2), 110–117. https://doi.org/10.1016/j.jarmac.2012.02.004.

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Lying about flying: The first experiment to detect false intent. *Psychology, Crime & Law, 17*(7), 611–620. https://doi.org/10.1080/10683160903418213.

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior, 33*(2), 159–166. https://doi.org/10.1007/s10979-008-9143-y.

Vrij, A., Leal, S., Mann, S. A., & Granhag, P. A. (2011). A comparison between lying about intentions and past activities: Verbal cues and detection accuracy. *Applied Cognitive Psychology, 25*(2), 212–218. https://doi.org/10.1002/acp.1665.

Zangrossi, A., Agosta, S., Cervesato, G., Tessarotto, F., & Sartori, G. (2015). "I didn't want to do it!" The detection of past intentions. *Frontiers in Human Neuroscience, 9,* 1–10. https://doi.org/10.3389/fnhum.2015.00608.

Zeigarnik, B. (1939). On finished and unfinished tasks. In E. D. Ellis (Ed.), *A source book of Gestalt psychology*. New York: Harcourt Brace. (Original work published in 1927).

Detecting Deceptive Intentions: Possibilities for Large-Scale Applications

Bennett Kleinberg, Arnoud Arntz and Bruno Verschuere

INTRODUCTION

In the 9/11 attacks, terrorists posed as regular passengers when they boarded and hijacked American Airlines Flight 11 (e.g., Wilgoren & Wong, 2001). What if one could have detected that they did not have the benign intention like other passengers of flying to San Francisco but that they instead had the malicious intent of committing a devastating terrorist attack. For law enforcement and intelligence practitioners, it is key to identify people with benign intent and those who need further security checks *before* they board an airplane. Terrorist attacks in New York, Madrid, London, and Brussels have motivated the academic deception community to develop methods that allow for the detection of deceptive intentions.

The vast majority of academic research on deception detection is limited to detecting deception on past events (Mac Giolla, Granhag, & Vrij, 2014; Vrij, Granhag, & Porter, 2010). However, as the 9/11 attacks illustrate, from a practitioner's perspective, it is the temporal dimension of the future that is of critical

B. Kleinberg (✉) · A. Arntz · B. Verschuere
University of Amsterdam, Amsterdam, The Netherlands
e-mail: bennett.kleinberg@ucl.ac.uk

A. Arntz
e-mail: A.R.Arntz@uva.nl

B. Verschuere
e-mail: B.J.Verschuere@uva.nl

B. Kleinberg
Department of Security and Crime Science, University College London, London, UK

importance when it comes to crime prevention, especially in the context of current threats of terrorist attacks. The aim of this chapter is to give an overview of the dominant deception detection theories and discuss existing interviewing approaches, methods, and cues to detecting deceptive intentions. We also outline which requirements an applied intentions-detection framework must meet.

Throughout this chapter, we focus on the applicability of existing approaches to real-world security processes on a large scale. We illustrate the challenges and requirements for applied deception detection tools on intentions through the example of airport passenger security operations. Throughout this chapter, we will adhere to Mac Giolla et al.'s (2014) definition of true and false intent. Accordingly, "true intent refers to a future action which [someone] *intends* to carry out, while […] false intent refers to a future action that [they do] *not* intend to carry out" (p. 155). Since false intent does not necessarily imply a criminal element, we define malicious intent as *a future action someone intends to carry out that causes harm to others*. Although from a researcher's perspective both the detection of false *and* malicious intent are worth investigating, it is mainly the malicious intent in which practitioners working in crime prevention are interested. For example, prospective passengers lying about flying to New York for a conference hiding that they are having an affair there (false intent) are less relevant than someone hiding that they are planning to carry out an attack (malicious intent). We will discuss and address this challenge in this chapter as well.

This chapter is structured as follows. First, we discuss the problem of low base rate settings and then define a set of *criteria* for the detection of deceptive intentions on a large scale paying particular attention to the requirements from an applied perspective. We will use these criteria to evaluate the dominant deception theories, interviewing approaches, methods, and deception cues in the next sections. Second, we give a brief overview of main *theories of deception*, namely arousal-based and cognition-based deception detection, and evaluate to what extent they may guide large-scale applications. Third, we examine which *interviewing approaches* are most useful for deception detection. Fourth, we discuss some *information elicitation methods* that may help to increase deception detection validity. Fifth, we discuss which *cues* are most applicable to the airport screening context.

The Paradox of the Low Base Rate in Applied Settings

For the course of this chapter, we define the following context to which the deception detection system (i.e., *the system*) could be applied. Consider the problem of airport security professionals who have to determine for vast numbers of passengers, whether they potentially have malicious intent with their trip or not. With change on its way toward more seamless passenger flows during the whole security process (i.e., ideally minimizing the number of security checks and making them as least intrusive as possible), an area of interest for practitioners is the *pre*-screening of passengers *before* they arrive at

the airport. Security processes at the airport could then flexibly be targeted at the specific intelligence requirements (e.g., which information needs further clarification) about each passenger. As such, a prospective screening system applied to that problem could function as the first filter in a system of multiple security layers, each of which would only subject those passengers to its test that "failed" the previous layers. By doing so, this system would address the problem of finding a needle in a haystack (e.g., someone with terrorist intentions among millions of ordinary passengers) by successively decreasing the size of the haystack. For the sake of argument, we will assume that the system will have to be able to screen up to 200,000 passengers each day on a single airport (e.g., London Heathrow: 205,000, Amsterdam Schiphol: 159,000, Paris Charles de Gaulle: 180,000; Airports Council International, 2016).

The large numbers of passengers, however, pose a particular statistical challenge for any screening tool. Let us assume a screening tool has a remarkable sensitivity (i.e., correctly identifying those that have malicious intent) and specificity (i.e., correctly identifying those that *do not* have malicious intent) of both 90%. What makes this particular context difficult, even for this highly accurate screening tool, is the low base rate (i.e., the small number of passengers with malicious intent). When the base rate is low (e.g., 0.0001, see Honts & Hartwig, 2014), even a highly accurate screening tool results in a large number of false positives; that is, it classifies ordinary passengers incorrectly as having malicious intent, simply as a function of the imbalance toward passengers without malicious intent. Table 21.1 illustrates that for 200,000 passengers, the percentage of correct identifications of malicious intentions when the screening tool indicates "*malicious intention*" (i.e., the precision of the screening tool) is effectively only 0.09%. Consequently, 99.91% of the cases when the screening tool signals "*malicious intention*" are false positives. This base rate paradox emphasizes the challenge of the passenger screening context and highlights the need for successive filters.

A cascading filter system could address this problem given that the assumption of statistical independence of indicators is met. Let us assume

Table 21.1 Illustration of the base rate problem for a fictitious screening tool with sensitivity and specificity of 90%

Outcome screening tool					
		Malicious intent	*No malicious intent*	*Total*	*Recall*
Reality	Malicious intent	18	2	20	**90%** (=sensitivity)
	No malicious intent	19,998	179,982	199,980	**90%** (=specificity)
	Total	20,016	179,984	200,000	
	Precision	**0.09%**	**99.99%**		

the screening tool A indicates the deceptive state X based on the criterion C_A with a sensitivity and specificity of 90% (Table 21.1). This, per definition, results in 10% false negatives and 10% false positives. Now let us further assume that the additional screening tool B is applied that also indicates deceptive state X but bases this decision on criterion C_B which is independent of (i.e., uncorrelated to) criterion C_A. In this case, the probabilities are conditional, that is, the percentages of miss-classifications[1] are multiplied and decrease to $0.10*0.10 = 0.01 = 1\%$. For n cascades, the probability of miss-classification is 0.01^n (e.g., for $n = 4$: $0.01^4 = 0.0001 = 0.01\%$). In the latter case, the precision of signaling "malicious intent" would increase from 0.09% ($n = 1$) to 50.00% ($n = 4$).

It is critical that cascades of security filters be independent of each other so that the occurrence of criterion C_B does not depend on the occurrence of criterion C_A; that is, both criteria are indicators of deceptive state X but measure it in different, unrelated ways. An example could be a system that indicates deception through the verbal content (e.g., what someone says) and the verbal style (e.g., how someone says it; see below). If both indicators are independent, this will allow for a combination of cascaded indicators through the analysis of verbal statements.

CRITERIA FOR LARGE-SCALE INTENTION DETECTION SYSTEMS

A deception detection system applicable within the context of prospective airport passenger screening also poses particular challenges from an applied perspective. In this part of the chapter, we describe which requirements—besides high accuracies of identifying passengers correctly—such a system must meet. We discuss specific elements of an applied large-scale deception detection system that refer to its applicability on real-life purposes such as prospective airport passenger screening. To grasp applicability in its full complexity, we briefly outline sub-criteria relevant to the applicability of deception detection systems on large scale (see Table 21.2 for a summary).

Large-Scale Data Collection

The process of gathering data useful for an assessment of whether a passenger is to be believed or not is referred to here as data collection. While many deception studies relied on face-to-face interviews after participants committed a mock crime (e.g., preparing to place a malware USB stick in a shopping mall; Sooniste, Granhag, Knieps, & Vrij, 2013), the applied context here precludes such procedures. Collecting data through face-to-face interviews is to date the most corroborated form of eliciting cues to deception. However, it is logistically not realistic to conduct interviews with all passengers at the airport or to perform any other kind of disruptive intervention in the natural flow of passengers such as hands-on psychophysiological measurements (but see

Table 21.2 Summary of applicability criteria for large-scale deception detection systems

Criterion	The deception detection system…	Research agenda
Large-scale data collection	… permits collecting statements/responses from vast numbers of airport passengers simultaneously	Which are techniques and methods most suitable for the screening of 200,000 passengers per day?
Real-time data analysis	… entails an instant, automated analytical process to derive veracity judgments	Can deception cues (and the veracity of statements) be assessed reliably in real time?
Implementability	… is practically and logistically fit to be used in existing passenger procedures	How can validated techniques be incorporated into existing airport procedures?
Customer friendliness	… does only require a minimal amount of time and effort from the passenger	Can prospective passenger screening be done in short time with little passenger effort?
Theory-based	… is based on scientific theory and has withstood scientific evaluation	Which techniques and methods are the most promising for the detection of deceptive intentions?
Flexibility	… can flexibly be adapted to security requirements	Which techniques and methods allow for the highest flexibility in veracity assessments?
Granularity	… can determine the veracity of units of analysis smaller than the whole statement (e.g. single utterances)	Can the analysis of statements be fine-tuned toward the detection of deception in single utterances?

Panasiti et al., 2016; Warmelink et al., 2011). That impediment suggests that alternative forms of data collection must be explored and adopted. For example, a more likely approach is to use existing procedures in the airport security process where passengers could be asked brief questions, such as the standard queuing for baggage screening or online check-in processes that are becoming the norm. Large-scale data collection implies that the system is scalable to scores of passengers. The scalability means that the deception detection method is suitable to be applied to a large number of travelers (e.g., 200k per day) and can relatively easily be *up*-scaled without extensive investments in human workforce and training.

Real-Time Data Analysis

The analytical process must be automated to derive near real-time veracity judgments. Standard procedure from interviewing studies is that participants deliver a verbal statement about, for example, their whereabouts during an alleged mock crime. That oral statement is then transcribed and handed over to one, or preferably more, independent human coders who score statements

on a range of criteria such as level of detail or plausibility (Sooniste et al., 2013; Sooniste, Granhag, Strömwall, & Vrij, 2015; Vrij, Granhag, Mann, & Leal, 2011; Warmelink, Vrij, Mann, & Granhag, 2013a). Manual coding is time-consuming and is currently not done in real time. Similarly, the number of tasks and checks performed by trained human coders—aside from the time constraints—is limited and prevents more complex tasks like verification of provided information and fine-grained coding of provided text statements (e.g., looking deeper than the overall text).

Implementability

A precondition for an approach to be used in real life is that it can be implemented into existing processes or by extending existing methods. For example, conducting face-to-face interviews with every passenger is not only logistically difficult, but it is also not implementable into the actual flow of current airport security systems because it takes too long and is too costly. In contrast to the general umbrella term of applicability, implementability has an additional, practical dimension, given logistical challenges, the feasibility of actually implementing a tool into security processes, as well as the potential of scaling the tool up to large numbers of people. From this follows that implementability subsumes applicability, but not vice versa.

Customer Friendliness

A further challenge for the application of deception detection tools is the inevitable compromise between academic rigor and stakeholders' interests. Although there are multiple aspects where the stakeholders' point of view might conflict with an academic's proposal (e.g., financial, ethical, theoretical), a noteworthy issue is the brevity of the developed system and the inconvenience caused to passengers. For external stakeholders, time is a premium and passengers' satisfaction is a vital ingredient for a thriving business. However, this puts the academic researcher into an unusual position. A standard polygraph examination, for instance, typically takes several hours (Meijer & Verschuere, 2010). Applied deception detection systems should ideally not exceed a few minutes' duration and should require as little effort from the passengers as possible (Honts & Hartwig, 2014). Computer-automated techniques would greatly facilitate data collection and veracity judgments within a short time.

Theory-Based

We think another requirement for a large-scale deception detection system is that it is built on a sound scientific theory. A simple "whatever works" approach is questionable for the airport screening context in the absence of

a guiding theory. Without a theoretical base for the tools used, any future development within that line of deception research will hang loosely in a vacuum of results without being able to derive predictions on how these results came about. With an increasing acceptance in psychological research of methods from machine learning, however, it will be interesting to see how large data-driven investigations compare to typically smaller, theory-led approaches (for an overview paper on the issue, see Yarkoni & Westfall, 2017).

To illustrate the need for scientific theory in the development of security tools, consider the extreme example of the IED detector called *Advanced Detection Equipment* (*ADE-101*). The *ADE-101* was sold to various governments with the promise that this device could "pick up the most minuscule traces of explosives, drugs, ivory and even money" (Morris, Jones, & Booth, 2013, para. 1). In fact, that device was little more than a golf ball finder sold by a fraudulent businessperson. The *ADE-101* had cost the Iraqi authorities alone more than GBP 55 million (Booth, 2013). Besides the obvious fraud involved in this case, there was no theory behind the alleged working mechanisms of the device, nor was there an empirical validation of its effectiveness.

Flexibility

A system applied for passenger screening purposes must be flexible on passenger numbers, security risk estimations, and specific flight characteristics. For example, when there is a heightened security risk (e.g., due to previous terrorist attacks), a large-scale screening system must be able to adapt to that situation by adjusting the cutoff used to make a decision. Dynamic filtering would imply altering the compromise between sensitivity (i.e., the true positives) and specificity (i.e., the true negatives). Under specific circumstances, specificity might be favored over higher sensitivity; under other circumstances, the opposite might be needed.

Granularity

Granularity refers to how fine-grained the judgments made by a deception detection tool are. That is, granularity represents a continuum from coarse judgments (e.g., liar vs. truth-teller) to finer resolutions such as single utterances. Whereas in controlled experimental studies, the liars are typically instructed to tell an outright lie (e.g., pretending to have played a game whereas, in fact, they stole money; Vrij, 2008), a lying passenger can likely *embed* their lie into a mainly honest account (see Mac Giolla et al., 2014). This implies that it does not longer suffice to use the person who is lying as the unit of analysis (i.e., who is a liar and who a truth-teller). Rather a more granular analysis is needed that permits the investigator to determine, ideally, *what* someone is lying *about*. As we will see later in this chapter, current verbal content-based cues perform relatively poor on this requirement, whereas

stylometric cues (e.g., Fornaciari & Poesio, 2013) may offer a path for the future. Alternatively, steps toward within-subjects deception investigations (i.e., having the same person tell a truthful account as well as a lie) might also offer a way to discern single deceptive aspects within whole statements (Vrij, 2016).

Deception Detection Theories

Most studies conducted on deception detection fall, broadly speaking, into one of the two dominant theories on deception: arousal-based versus cognition-based deception detection. They are rooted in different assumptions about the relationship between the mental state of deception and the cues through which this mental state is detectable.

Arousal-Based Deception Detection

The arousal theory holds that the mental state of lying can be inferred from arousal associated with lying (Vrij et al., 2010). The arousal-based assumption holds that the involuntary display of physiological signs of arousal is informative to the mental state of deception. For example, research into micro-expressions (Ekman, 2009; Schubert, 2006) poses that minimal muscular activity in the facial area is a cue to deception. Likewise, methods such as the Screening Passengers by Observation Technique(s) (SPOT; Honts & Hartwig, 2014) assume that lying is uniquely related to physiological and behavioral signs including body language and micro-expressions (see Honts & Hartwig, 2014; Perry & Gilbey, 2011). Consequently, a liar would be detectable through the mere observation of their overt behavior (Panasiti et al., 2016; Warmelink et al., 2011).

Cognition-Based Deception Detection

Starting with the notion that lying is cognitively more demanding than telling the truth (e.g., Zuckerman, DePaulo, & Rosenthal, 1981), the rationale of cognition-based deception detection is that increased cognitive load that comes along with lying results in the leakage of cues to deception (Masip, Sporer, Garrido, & Herrero, 2005; Oberlader et al., 2016; Vrij & Granhag, 2012). Various aspects reasonably make the act of lying harder than telling the truth. Consider a passenger flying to New York City (NYC) for a geography conference. When interviewed about the conference, they can easily tell about their plans and the preparation involved in their trip. Now let us assume a terrorist is planning an attack on NYC, but who claims to fly to NYC for the geography conference. The liar's task of convincing the investigator is probably harder than that for the truth-teller. Not only can the liar be thought of operating two accounts of their trip simultaneously, but they also

have to maintain a convincing false account without risking the leakage of any hints alluding to their malicious plans.

Evaluation

The primary concern about arousal theory is that it falls prey to the so-called Othello error (for a historical explanation, see Vrij et al., 2010). The Othello error means that one ignores alternative explanations for the display of alleged cues to deception. While signs of stress may well accompany someone's lying, this does not exclude the possibility that someone telling the truth shows the same signs of stress. In a context such as an airport passenger screening, the issue of innocent stress becomes evident when one realizes that passengers rushing to catch their flight or traveling with small children will display physiological signs similar to those that are postulated to be uniquely related to lying.

There is increasing support for the cognition-based deception theory (Meissner et al., 2014; Vrij, Fisher, & Blank, 2017; but see also the critique by Levine, Blair, & Carpenter, 2018; and the response by Vrij, Blank, & Fisher, 2018). Despite the substantial evidence, it merits attention that the Othello error could also be at play for cognition-based deception theory. First, the rationale that lying is harder than telling the truth might not hold true for well-prepared or repeated lies. For example, if someone is repeatedly flying to NYC under pretense, the false story (e.g., a conference covering for an affair) is rehearsed and might therefore not be more difficult to tell than the truthful story. Second, the relative ease of telling the truth also depends on the complexity of the truth. Someone telling a complex true story (e.g., that they meet at a secret government facility for a classified meeting) may find it difficult to appear convincing—similar to someone lying about an activity. If someone is flying to a secret meeting about which they are not supposed to talk, the truthful account might be more difficult to tell than a simpler false account (i.e., that they are flying to a conference). Although the Othello error is arguably less problematic for cognition-based deception theory, future research should address these questions to refine cognition-based deception theory further.

Based upon the scientific support, the cognitive deception theory seems more promising than arousal-based deception detection in the context of passenger screening. Regardless of the particular deception theory, cues to deception (i.e., nonverbal and verbal indicators of the interviewee that are informative to the veracity of the statement made by the interviewee) are small and unreliable (DePaulo et al., 2003). The use of those cues for deception detection, therefore, requires approaches that can increase the occurrence of the cues in truth-tellers and decrease the occurrence in liars, or vice versa (Vrij & Granhag, 2012). In the next part of this chapter, we discuss approaches to eliciting cues to deception.

<center>APPROACHES TO SUSPECT INTERVIEWING</center>

No matter how brief (e.g., "No") or long (e.g., "I booked my ticket last Thursday online through..."), the minimal requirement for any deception detection approach is a statement. There are two broad approaches to eliciting a statement from an interviewee (Meissner et al., 2014): the accusatorial approach and the information-gathering approach.

The Accusatorial Approach

The accusatorial approach to suspect interviewing is based on the rationale that the interviewer needs to engage actively in the interview to elicit admissions of intentional wrongdoing (Meissner et al., 2014). The accusatorial approach involves an interviewer who is trained to exert control over the interviewee, applies techniques to manipulate the interviewee psychologically, and typically asks closed (e.g., yes/no) questions. A formulation of the widely used accusatorial approach is the Reid technique (Gallini, 2010; Kassin et al., 2010), which consists of two phases. In the first step, the suspect is interviewed to determine whether they are indeed a likely suspect. The second phase of the Reid technique consists of techniques targeted at obtaining confessions from the suspect through a set of techniques that manipulate the suspect (i.e., custody and isolation, confrontation, minimization; see Kassin and Gudjonsson, 2004).

The Information-Gathering Approach

According to Meissner et al. (2014; see also Swanner, Meissner, Atkinson, & Dianiska, 2016), the key ingredients of the information-gathering approach are establishing rapport with the interviewee (e.g., positive affirmations, interest, calmness; see Evans et al., 2014), using positive confrontation and asking open-ended questions. The goal of the information-gathering approach—eliciting information—is strikingly different from the accusatorial approach which is obtaining confessions. Rooted in the cognitive interview (e.g., Fisher & Geiselman, 1992; Fisher, Geiselman, & Amador, 1989; for a recent meta-analytical overview, see Memon, Meissner, & Fraser, 2010), the information-gathering approach has a clear focus on treating the interviewee as a source of information rather than the possessor of guilt. As a result, the interviewing of victims, witnesses, and suspects does not radically differ since the goals are always to obtain relevant information.

Evaluation

Opponents of the accusatorial approach have voiced concerns about the fairness of the approach toward innocent interviewees. Studies suggest that the accusatorial approach does elicit confessions in guilty suspects but fails to

protect those who have not committed any crime. Innocent subjects were found to provide false confessions merely as a function of a coercive interviewing style (Loney & Cutler, 2016; Meissner et al., 2014). Meissner et al. (2014) found the information-gathering approach better able to elicit true confessions and reduced the rate of false confessions as compared to the accusatorial approach. Not only does the elicitation of false confessions conflict with ethical standards in most countries (e.g., Soukara, Bull, Vrij, Turner, & Cherryman, 2009), they also impede the validity of the investigative interview (i.e., they do not elicit useful information). The accusatorial interviewing approaches (e.g., the Reid technique, see Kassin et al., 2010; the Behavioral Analysis Interview, Inbau, 2013) are the standard interviewing procedure in the US but have been banned from European countries as well as from the British, New Zealand, and Australian judicial system (Kassin et al., 2010; Meissner et al., 2014). Looking ahead, the release of the FBI's High-Value Detainee Interrogation Group (HIG) report (High-Value Detainee Interrogation Group, 2016) suggests that the US is also moving toward actively advocating an information-gathering approach. It is the information-gathering approach that has become the standard approach used in academic deception research (Vrij et al., 2017) and has been proposed for law enforcement and intelligence applications (Granhag, Vrij, & Meissner, 2014).

Since the accusatorial approach works mainly for the elicitation of confessions, one can imagine how this conflicts with the applicability. Given the low base rate of airport passengers with malicious intent, an approach that is inherently biased toward false positives such as the accusatorial approach, will lead to unsatisfyingly large numbers of false alarms. An unnecessarily inflated large number of false positives is highly undesirable from both a security practitioner's point of view since it redirects resources away from the actual problem (i.e., finding the true positives) and from the airport authority's perspective since each false alarm implies a falsely accused customer.

The accusatorial approach inherently assumes guilt, making an interview resemble an interrogation and putting interviewees directly under suspicion. To the contrary, the information-gathering approach is embedded in conservative assumptions about information that truth-tellers can provide but liars cannot, which has been shown to yield both more true confessions and fewer false confessions than the accusatorial approach. The information-gathering approach is more applicable to low base rate settings and is less offensive toward airport passengers of which the vast majority (e.g., 99.999%) has no malicious intent. Despite the moderate accuracy rates of the information-gathering approach, its conservative assumptions about the relationship between behavior and deception make it a more suitable approach for prospective airport passenger screening than the accusatorial approach.

Sooniste et al. (2015) found support for the information-gathering interviewing approach for the detection of true and false intent. Half of the participants planned a mock crime (i.e., installing malware on a university computer), whereas the other half planned an innocent activity

(i.e., organizing a protest). Before enacting their task, both groups were instructed to convince an interviewer that they were organizing the protest. When participants were interviewed with the information-gathering approach (i.e., using the cognitive interview by establishing rapport, reinstating memory, and encouraging rich descriptions of activities) compared to a standard interviewing technique (i.e., without any of the information-gathering techniques), both liars and truth-tellers provided more detailed information. Crucially, the amount of information provided in the information-gathering interviews allowed for better truth–lie discrimination.

Based on the available scientific support and the higher customer friendliness, we believe the information-gathering interviewing approach is more suitable for airport passenger screening than the accusatorial approach. In the next part, we outline and evaluate different methods to eliciting information used within the information-gathering approach.

Methods for Eliciting Information

Within the information-gathering interviewing approach, several specific methods have been used to obtain more diagnostic veracity information from interviewees. In this part, we outline three promising methods relevant to the context of airport passenger screening.

Imposing Cognitive Load

An important method of cognition-based deception detection is imposing additional cognitive load to make the interview situation more cognitively demanding (Vrij et al., 2017). In particular, building on the existing differences in the cognitive effort involved in telling the truth versus lying, actively imposing additional cognitive load is postulated to create a situation that is even more difficult for the liar than for the truth-teller. Note, however, that imposing cognitive load is only one method of the cognitive approach to deception detection and these terms should not be used interchangeably (Vrij & Fisher, 2016). The rationale is that by directing mental efforts to a secondary task (e.g., maintaining eye contact, Vrij et al., 2010; holding a weight, Debey, Verschuere, & Crombez, 2012), cognitive resources become scarcer for the liar. Without many cognitive resources left, the liar will find it even harder to maintain a convincing false story; that is, they will have more trouble to lie successfully. Similarly, Vrij et al. (2008) proposed to instruct interviewees to recall an event in the reverse order. While this should be easier for truthtellers, liars will be confronted with heightened cognitive load (see below).

Asking Unanticipated Questions

Based on the assumption that liars prepare for a suspect interview, Vrij et al. (2009) reasoned that providing spontaneous stories would be harder for

liars than for truth-tellers (see also DePaulo et al., 2003; Masip et al., 2005). Whereas the liar and the truth-teller would be able to provide convincing answers to those questions that they expect, only the truth-teller will be able to do so for less expected questions where an answer needs to be formulated on the spot.

The Model Statement Technique

Recently, the information-gathering approach has been extended by providing an example of a detailed answer (e.g., Harvey, Vrij, Nahari, & Ludwig, 2017; Leal, Vrij, Warmelink, Vernham, & Fisher, 2015). The idea behind the so-called model statement technique is that although interviewees are typically asked to provide highly detailed answers, it is not certain whether they are aware of exactly how detailed that answer must be. One way to help interviewees is to provide them with an example of a detailed account of an event containing the expected level of detail.

Evaluation

Imposing cognitive load: Vrij et al. (2008; see also Evans, Michael, Meissner, & Brandon, 2013) imposed additional load on interviewees by asking half of their participants to recall an event in the regular, chronological order (i.e., beginning with the most distant), whereas the other half was instructed to remember the event in the reverse order (i.e., starting with the most recent and going stepwise back in time). Differences between truth-tellers and liars were magnified in the reverse order technique and allowed for better discrimination (for conflicting evidence, see Fenn, McGuire, Langben, & Blandón-Gitlin, 2015).

Asking unanticipated questions: In a first experiment, Vrij et al. (2009) devised an interview that asked a set of both anticipated questions (e.g., "Can you tell me in as much detail as possible what you did in the restaurant?") and *un*anticipated questions (e.g., "In relation to the front door, where did you sit?"; Vrij et al., 2009, p. 162) about the truthful or deceptive accounts of participants' whereabouts in a restaurant. Unanticipated questions revealed larger truth–lie differences than expected questions, especially if the unanticipated questions were about the spatial arrangement of the restaurant. The unexpected questions method has been used in multiple studies successfully (e.g., Shaw et al., 2013; Warmelink, Vrij, Mann, Leal, & Poletiek, 2013) and has emerged as a valuable method for exploiting and increasing differences between truth-tellers and liars (Vrij et al., 2017).

To test the unanticipated question technique on deceptive intentions, Sooniste et al. (2013) gave participants either an innocuous mission (i.e., buying gifts) or a mock criminal mission (i.e., placing a malicious USB stick in a shopping mall). Both groups prepared for this task, but those with a

mock criminal story also developed their cover story, which was conceptually the same as the innocuous task. The authors found that questions about the planned activities directly did not elicit truth–lie differences. However, when asking questions regarding the planning of the activity (e.g., "What was the main goal of your planning?"), the truthful answers were perceived as more detailed than lies. A potential implication of this finding is that liars plan the activity in a different way than truth-tellers do. Although not many studies have investigated this explicitly yet, a possible explanation could stem from the questions anticipated by the liar: Sooniste et al.'s findings suggest that liars were more prepared to answer intentions-related questions than planning questions.

The model statement technique: Leal et al. (2015) presented participants with an audio-taped statement about an event (e.g., a Formula 2 race) unrelated to the research scenario (e.g., false or genuine insurance claims). They found that receiving the model statement previous to giving the statement affects truth-tellers and liars in different ways. For liars, there was no change in the human-judged plausibility between those who did and did not receive the model statement. To the contrary, for truth-tellers, the model statement resulted in more plausible statements, which suggests that the model statement was beneficial to the overall classification accuracy (non-cross-validated accuracies: 62.5% vs. 80.0%, without and with the model statement, respectively; for null-findings regarding the model statement see Bogaard, Meijer, & Vrij, 2014; Brackmann, Otgaar, Roos af Hjelmsäter, & Sauerland, 2017). To date, the model statement technique has not been assessed for the detectability of deceptive intentions.

Of the information elicitation methods discussed, the model statement method and asking unanticipated questions are the most promising for the detection of deceptive intentions (see also Vrij & Fisher, 2016). Imposing cognitive load is less applicable to the context of prospective passenger screening since it often requires active engagement with a secondary task or is related to future events that have not yet happened (e.g., for the reverse order technique). While the unanticipated questions method has successfully been used in experimental studies on intentions, for the model statement technique future research will have to explore how well they are suited for the study of deceptive intentions. Box 21.1 highlights important challenges for the research agenda of the detection of deceptive intentions. We will next discuss verbal cues that are used to detect deception.

VERBAL CUES TO DECEPTION

Hundreds of cues have been proposed to determine whether a suspect is answering truthfully or deceptively (DePaulo et al., 2003). Cues range from behavioral (e.g., head nods, fidgeting) and physiological (e.g., eye muscle

Box 21.1 Outlook on the research agenda for the detection of deceptive intentions on a large scale

---

Research agenda

–Does model statement technique facilitate detection of deceptive intentions?
–Is the verifiability of detail rationale applicable to the detection of deceptive intentions?
–Can the scoring of verbal cues be computer-automated?
–Can the information-gathering approach be automated and shortened (e.g., chat-based information elicitation)?
–Can stylometric analysis be used to determine the content of lies?
–Can two (or more) independent (i.e., uncorrelated) verbal deception cues be derived from verbal statements (for cascaded screening)?

---

movements) to speech-related (e.g., vocal tension, pitch) and content-based cues (e.g., spontaneous corrections). From the perspective of the implementability, large-scale data collection, and granularity, we focus on verbal cues to deception, whereby we differentiate between content-based cues, the verifiability of information, and stylometric cues.

Content-Based Cues

Verbal deception detection assumes that the content of a statement (i.e., *what* the suspect says) is informative to the veracity of the declaration. Reality Monitoring provides a theoretical backcloth as to why the content of deceptive versus truthful statements would differ. Originally developed by Johnson and Raye (1981; Johnson, Bush, & Mitchell, 1998; Nahari & Vrij, 2014), Reality Monitoring was used to identify the source of a memory of an event. According to Reality Monitoring, a memory can be attributed either to an external source or to an internal source. A memory originating from the internal source has been constructed through cognitive operations (i.e., forming a memory of how the event *could have been*), whereas a memory attributable to the external source has been obtained through perceptual processes (i.e., the event *has been experienced genuinely*). The verbal accounts of events would, therefore, represent the source of the corresponding memory. If a memory stems from the external source, the account of the event in question should be richer in temporal, spatial, and perceptual details and should be more realistic, "reconstructable," and richer in affect than accounts of memories from the internal (i.e., fantasized) source.

The Verifiability Approach

An important addition to verbal cues to deception originates from the Verifiability Approach (Nahari, Leal, Vrij, Warmelink, & Vernham, 2014; Nahari, Vrij, & Fisher, 2014a). Liars face the dilemma between providing a believable

account with sufficient detail and, at the same time, not mentioning any poten-
tially incriminating information (i.e., those details that the interviewer could
verify). Research showed that liars evade this dilemma by providing unverifia-
ble details (Nahari et al., 2014a). For example, "I booked the trip together with
someone I know" contains some details but is mainly non-verifiable, whereas "I
booked the flight to New York with my friend Paul last Thursday" adds verifiable
context to the same proposition. The Verifiability Approach set out to exploit
this strategy by looking at how many verifiable details true and false statements
contain.

Stylometric Cues

Rather than looking at *what* people convey in their verbal reports, research-
ers have also attempted to differentiate truthful from deceptive statements
through *how* people convey their stories. The technique of stylometry pos-
tulates that one can attribute the identity of the author of a text to the sty-
listic features used in the text (e.g., Fornaciari & Poesio, 2013; Luyckx &
Daelemans, 2008). In a stylometric analysis, a text is decomposed into
features pertaining to *how* the text is written rather than which content the
text conveys. Within stylometric analysis, Schler, Koppel, Argamon, and
Pennebaker (2006) distinguish between surface-related features (e.g., the use
of grammatical function words) and content-related features (i.e., the mean-
ing of the words). In contrast to verbal content cues, the content-related
features in stylometric analysis often stem from lexicons (e.g., the Linguistic
Inquiry and Word Count software, Pennebaker, Boyd, Jordan, & Blackburn,
2015), assigning each word to a psychological dimension, for example.
Whereas verbal content cues are about the semantic qualities of a whole state-
ment, in the stylometric and linguistic analysis, the content-related features
often are about the meaning and/or function of single words or tokens.
Researchers typically use techniques from supervised machine learning to
build algorithmic classifications of truthful and deceptive texts using a num-
ber of stylometric features (e.g., Fornaciari & Poesio, 2013, 2014; Mihalcea
& Strapparava, 2009; Ott, Cardie, & Hancock, 2013; Ott, Choi, Cardie, &
Hancock, 2011).

Evaluation

Content-based cues and Reality Monitoring: Using verbal content-based
cues for the detection of deceptive intentions has only occurred since recently
(e.g., Kleinberg, Nahari, Arntz, & Verschuere, 2017; Vrij, Granhag, et al.,
2011; Warmelink, Vrij, Mann, & Granhag, 2013b). Vrij, Granhag, et al.
(2011) conducted the first study using information-gathering interviewing
principles to detect lies about intentions. In their experiment, they instructed
departing passengers at an airport to either tell the truth about their upcom-
ing flight or lie about it. In a subsequent interview, each participant answered

a set of questions that were then transcribed and coded by human judges on content-based variables. They found that truth-tellers' statements were more plausible than liars' statements, contained more complications and more spontaneous corrections. Moreover, in another experiment, researchers compared the level of detail and plausibility of true and false statements about both past events and intentions (Vrij, Leal, Mann, & Granhag, 2011). They found that truthful answers to intention-related questions were more detailed and more plausible than deceptive answers.

Masip et al. (2005) found that the Reality Monitoring verbal content analysis tool is useful for the discrimination between truthful and deceptive statements with classification accuracy rates ranging between 65 and 75%. Separate cues from Reality Monitoring that have been shown to differ between deceptive and truthful statements are especially the plausibility of a statement (e.g., Leal et al., 2015; Vrij, Granhag, et al., 2011) and the richness of detail (e.g., Vrij et al., 2008; Warmelink, Vrij, Mann, Jundi, & Granhag, 2012).

The Verifiability Approach: A series of studies (e.g., Harvey et al., 2017; Jupe, Leal, Vrij, & Nahari, 2017; Nahari et al., 2014a) found that by looking at the amount of verifiable details, the discriminatory accuracy of verbal content analysis can be increased further with accuracy rates ranging between 67 and 90%. It is noteworthy that the Verifiability Approach seems relatively robust against countermeasures. When liars and truth-tellers were informed that verifiable details are indicative of the truthfulness of a statement, truth-tellers but not liars were able to provide more verifiable details (Nahari, Vrij, & Fisher, 2014b). Liars might simply not be willing to risk providing highly detailed information that the interviewer could potentially use against them (see also Kleinberg, Nahari, & Verschuere, 2016). Research on the Verifiability Approach for the detection of false intent is emerging and seems a worthwhile avenue for future research (Jupe et al., 2017).

Stylometric cues: With advances in computational methods, stylometric analysis has become more widespread in deception research (Fitzpatrick, Bachenko, & Fornaciari, 2015). For example, Ott et al. (2013) used machine learning classifiers to predict whether hotel reviews were truthful or deceptive. By adding variables such as the use of self-references (e.g., personal pronouns) and use of negative affect in the hotel reviews, they were able to devise a classifier that achieved up to 89.3% accuracy (see also Mihalcea & Strapparava, 2009; Ott et al., 2011). Using the same hotel review dataset, Feng and Hirst (2013) built stylometric profiles of deceptive and truthful hotel reviews (i.e., an average of a deceptive/truthful hotel review) and compared the profile compatibility of each review with the mean profile. They obtained a classification accuracy of up to 90.1%. Recent findings suggest that a combination of methodologies from computational linguistics (e.g., lexicon approaches and named entity recognition) might be a fruitful way to synthesize verbal deception theory and automated classification approaches (Kleinberg, Mozes, Arntz, & Verschuere, 2018).

Granularity

Content-based verbal cues such as the richness of detail are relatively ill-equipped to identify the veracity of smaller units of analysis (e.g., sentences, propositions, or utterances), but evidence suggests that stylometric cues might be useful to obtain a more granular level of analysis. For example, by zooming in on smaller parts within the entire statement, researchers changed the unit of analysis from whole texts to unique propositions made in court statements (Bachenko, Fitzpatrick, & Schonwetter, 2008; see also Fornaciari & Poesio, 2013). When the authors modeled verbal content constructs like inconsistencies using different indicators (e.g., verbal hedges: "maybe," "I guess"; verb tense change, thematic role change, noun phrase changes), they were able to correctly identify 75.6 and 73.8% of false and true propositions, respectively. Similarly, the verifiability of details might also offer paths toward more granular analyses: If the verifiability is used as a test of deception, rather than obtaining overall counts of verifiable details for a whole statement, one could explore whether small bits of verifiable information (e.g., names in single utterances) are informative to the truthfulness of parts of the entire declaration.

All of the three classes of verbal cues seem promising for the context of airport passenger screening. The more granular stylometric analysis and the paths open for the verifiability of details make these two types of cues particularly promising. Box 21.1 highlights questions for the research agenda on deceptive intentions.

The Controlled Cognitive Engagement

The Controlled Cognitive Engagement (CCE; Ormerod & Dando, 2015) is an illustration of a promising system that incorporates several of the discussed elements. To date, the CCE is the most extensive investigation of cognition-based deception detection on a larger scale in an airport screening context. The authors formulated six cornerstones for the CCE. The CCE method (1) was built on strategic interviewing principles, (2) aimed to elicit rich verbal accounts, (3) included tests of expected knowledge (e.g., someone claiming to fly to NYC should know where they are staying), (4) restricted verbal maneuvering (i.e., that the interviewee takes over the conversation), (5) contained elements that raised the cognitive load of respondents, and (6) looked at the content of statements to assess their veracity. Key features of the developed CCE method were question cycles consisting of an open question (e.g., "Please tell me about your plans in New York." [answer: "I'm attending the Geology conference there."]), a related focus question (e.g., "Who will you meet at the conference?" [answer: "Paul Johnson"]), and a test question (e.g., "Where do you know Paul Johnson from?" [answer: "He was my dissertation supervisor"]). By formulating this structure of asking questions without

specifying the exact questions to be asked, the CCE method is sufficiently flexible to allow for custom-made interviews depending on each passenger's context. After training airport security practitioners in the use of CCE, Ormerod and Dando (2015) compared how well CCE-trained officers were able to identify participants who tried to pass through airport security with a fake identity. The CCE method (66% of mock passengers identified) outperformed the widely adopted yet not scientifically corroborated suspicious signs method (3%, i.e., identifying passengers based on their physical display of suspicious behavior). Despite its successful test in the reported experiment, the CCE method has yet to replicate independently. Moreover, further research must establish how systems similar to the CCE can be useful for prospective passenger screening while meeting the specific applied requirements formulated in this paper (e.g., for 200k passengers in a fast, non-intrusive way). In particular, the issues of scalability and the requirement to screen passengers before they arrive at the airport merit special attention. Nevertheless, the CCE illustrates how theory-based techniques can be used for airport passenger screening purposes, and it is imaginable that a system for prospective passenger screening is combined with in situ CCE screening in a cascaded system. Future research will have to extend such techniques toward even shorter, possibly non-intrusive methods (e.g., chat-based information elicitation online).

Conclusions

This chapter set out to review deception detection research for the applied context of prospective airport passenger screening. As a guideline for the various research aspects (theories, interviewing approaches, methods to information elicitation, and verbal cues to deception), we defined a set of requirements of an applicable deception detection system. The cognition-based deception theory and the information-gathering approach seem most promising. Both were found to be more supported by evidence and to fit the applied requirements better. Furthermore, asking unanticipated questions and the model statement technique are promising methods for the elicitation of useful information. Lastly, it was found that three kinds of verbal deception cues are relevant for the applied context (i.e., content-based cues, the verifiability of details, and stylometric cues), with the verifiability of details and the stylometric cues to be the most promising. This chapter closed with an illustration of the Controlled Cognitive Engagement as a potential predecessor tool for a truly large-scale prospective airport passenger screening tool.

Note

1. Here, miss-classification refers to both false positives and false negatives given that we assume a sensitivity and specificity of both 90%.

<center>REFERENCES</center>

Airports Council International. (2016). Year to date passenger traffic, December 2015. Retrieved from http://www.aci.aero/Data-Centre/Monthly-Traffic-Data/Passenger-Summary/Year-to-date.

Bachenko, J., Fitzpatrick, E., & Schonwetter, M. (2008). Verification and implementation of language-based deception indicators in civil and criminal narratives. In *Proceedings of the 22nd International Conference on Computational Linguistics— Volume 1* (pp. 41–48). Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1599087.

Bogaard, G., Meijer, E. H., & Vrij, A. (2014). Using an example statement increases information but does not increase accuracy of CBCA, RM, and SCAN: Using an example statement with truth tellers and liars. *Journal of Investigative Psychology and Offender Profiling, 11*(2), 151–163. https://doi.org/10.1002/jip.1409.

Booth, R. (2013). Fake bomb detector conman jailed for 10 years. *The Guardian*. Retrieved from https://www.theguardian.com/uk/2013/may/02/fake-bomb-detector-conman-jailed.

Brackmann, N., Otgaar, H., Roos af Hjelmsäter, E., & Sauerland, M. (2017). Testing a new approach to improve recall in different ages: Providing witnesses with a model statement. *Translational Issues in Psychological Science*, *3*(2), 131–142. https://doi.org/10.1037/tps0000116.

Debey, E., Verschuere, B., & Crombez, G. (2012). Lying and executive control: An experimental investigation using ego depletion and goal neglect. *Acta Psychologica, 140*(2), 133–141. https://doi.org/10.1016/j.actpsy.2012.03.004.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (4th ed.). New York: W. W. Norton.

Evans, J. R., Houston, K. A., Meissner, C. A., Ross, A. B., LaBianca, J. R., Woestehoff, S. A., & Kleinman, S. M. (2014). An Empirical Evaluation of Intelligence-gathering Interrogation Techniques from the United States Army Field Manual: Intelligence-gathering interrogation techniques. *Applied Cognitive Psychology*, *28*(6), 867–875. https://doi.org/10.1002/acp.3065.

Evans, J. R., Michael, S. W., Meissner, C. A., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a psychologically based credibility assessment tool. *Journal of Applied Research in Memory and Cognition, 2*(1), 33–41. https://doi.org/10.1016/j.jarmac.2013.02.002.

Feng, V. W., & Hirst, G. (2013). Detecting deceptive opinions with profile compatibility. In *IJCNLP* (pp. 338–346). Retrieved from ftp://128.100.3.31/dist/gh/Feng+Hirst-IJCNLP-2013.pdf.

Fenn, E., McGuire, M., Langben, S., & Blandón-Gitlin, I. (2015). A reverse order interview does not aid deception detection regarding intentions. *Frontiers in Psychology*, *6*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4553365/.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Thomas.

Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology, 74*(5), 722–727. https://doi.org/10.1037/0021-9010.74.5.722.

Fitzpatrick, E., Bachenko, J., & Fornaciari, T. (2015). *Automatic detection of verbal deception* (Vol. 8). Morgan & Claypool Publishers. Retrieved from http://www.morganclaypool.com/doi/abs/10.2200/S00656ED1V01Y201507HLT029.

Fornaciari, T., & Poesio, M. (2013). Automatic deception detection in Italian court cases. *Artificial Intelligence and Law, 21*(3), 303–340. https://doi.org/10.1007/s10506-013-9140-4.

Fornaciari, T., & Poesio, M. (2014). *Identifying fake Amazon reviews as learning from crowds.* Association for Computational Linguistics. Retrieved from http://repository.essex.ac.uk/id/eprint/14591.

Gallini, B. (2010). Police "science" in the interrogation room: Seventy years of pseudo-psychological interrogation methods to obtain inadmissible confessions. *Hastings Law Journal, 61*, 529–577. https://scholarworks.uark.edu/lawpub/29/.

Granhag, P. A., Vrij, A., & Meissner, C. A. (2014). Information gathering in law enforcement and intelligence settings: Advancing theory and practice: Effective human intelligence gathering techniques. *Applied Cognitive Psychology, 28*(6), 815–816. https://doi.org/10.1002/acp.3093.

Harvey, A. C., Vrij, A., Nahari, G., & Ludwig, K. (2017). Applying the verifiability approach to insurance claims settings: Exploring the effect of the information protocol. *Legal and Criminological Psychology, 22*(1), 47–59.

High-Value Detainee Interrogation Group. (2016). Interrogation best practices report. Federal Bureau of Investigation. Retrieved from https://www.fbi.gov/file-repository/hig-report-august-2016.pdf/view.

Honts, C., & Hartwig, M. (2014). Credibility assessment at portals. In D. C. Raskin, C. Honts, & J. Kircher (Eds.), *Credibility assessment: Scientific research and applications* (pp. 37–62). San Diego: Academic Press.

Inbau, F. E. (Ed.). (2013). *Criminal interrogation and confessions* (5th ed.). Burlington, MA: Jones & Bartlett Learning.

Johnson, M. K., Bush, J. G., & Mitchell, K. J. (1998). Interpersonal reality monitoring: Judging the sources of other people's memories. *Social Cognition, 16*(2), 199–224.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67–85. http://dx.doi.org/10.1037/0033-295X.88.1.67.

Jupe, L. M., Leal, S., Vrij, A., & Nahari, G. (2017). Applying the verifiability approach in an international airport setting. *Psychology, Crime & Law, 23*(8), 812–825. https://doi.org/10.1080/1068316X.2017.1327584.

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions, risk factors, and recommendations: Looking ahead. *Law and Human Behavior, 34*(1), 49–52. https://doi.org/10.1007/s10979-010-9217-5.

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest, 5*(2), 33–67. https://doi.org/10.1111/j.1529-1006.2004.00016.x.

Kleinberg, B., Mozes, M., Arntz, A., & Verschuere, B. (2018). Using named entities for computer-automated verbal deception detection. *Journal of Forensic Sciences, 63*(3), 714–723. https://doi.org/10.1111/1556-4029.13645.

Nahari, G., Leal, S., Vrij, A., Warmelink, L., & Vernham, Z. (2014). Did somebody see it? Applying the verifiability approach to insurance claim interviews: The verifiability approach in insurance interviews. *Journal of Investigative Psychology and Offender Profiling, 11*(3), 237–243. https://doi.org/10.1002/jip.1417.

Nahari, G., & Vrij, A. (2014). Are you as good as me at telling a story? Individual differences in interpersonal reality monitoring. *Psychology, Crime & Law, 20*(6), 573–583.

Nahari, G., Vrij, A., & Fisher, R. P. (2014a). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology, 19*(2), 227–239. https://doi.org/10.1111/j.2044-8333.2012.02069.x.

Nahari, G., Vrij, A., & Fisher, R. P. (2014b). The verifiability approach: Countermeasures Facilitate its ability to discriminate between truths and lies: The verifiability approach and countermeasures. *Applied Cognitive Psychology, 28*(1), 122–128. https://doi.org/10.1002/acp.2974.

Oberlader, V. A., Naefgen, C., Koppehele-Goesel, J., Quinten, L., Banse, R., & Schmidt, A. F. (2016). Validity of content-based techniques to distinguish true and fabricated statements: A meta-analysis. *Law and Human Behavior, 40*(4), 440–457.

Ormerod, T. C., & Dando, C. J. (2015). Finding a needle in a haystack: Toward a psychologically informed method for aviation security screening. *Journal of Experimental Psychology: General, 144*(1), 76–84. https://doi.org/10.1037/xge0000030.

Ott, M., Cardie, C., & Hancock, J. T. (2013). Negative deceptive opinion spam. In *HLT-NAACL* (pp. 497–501). Retrieved from http://www.aclweb.org/website/old_anthology/N/N13/N13-1.pdf#page=535.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies—Volume 1* (pp. 309–319). Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=2002512.

Panasiti, M. S., Cardone, D., Pavone, E. F., Mancini, A., Merla, A., & Aglioti, S. M. (2016). Thermal signatures of voluntary deception in ecological conditions. *Scientific Reports, 6*(1). https://doi.org/10.1038/srep35174.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015.* Retrieved from https://repositories.lib.utexas.edu/handle/2152/31333.

Perry, M., & Gilbey, A. (2011). The screening of passengers by observation techniques programme. *Aviation Security International, 17*(3), 12.

Schler, J., Koppel, M., Argamon, S., & Pennebaker, J. W. (2006). Effects of age and gender on blogging. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs* (Vol. 6, pp. 199–205).

Schubert, S. (2006). A look tells all. *Scientific American.* Retrieved from http://www.scientificamerican.com/article.cfm?id=a-look-tells-all.

Shaw, D. J., Vrij, A., Leal, S., Mann, S., Hillman, J., Granhag, P. A., et al. (2013). Expect the unexpected? Variations in question type elicit cues to deception in joint interviewer contexts. *Applied Cognitive Psychology, 27*(3), 336–343.

Sooniste, T., Granhag, P. A., Knieps, M., & Vrij, A. (2013). True and false intentions: Asking about the past to detect lies about the future. *Psychology, Crime & Law, 19*(8), 673–685. https://doi.org/10.1080/1068316X.2013.793333.

Sooniste, T., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2015). Statements about true and false intentions: Using the cognitive interview to magnify the differences. *Scandinavian Journal of Psychology, 56*(4), 371–378. https://doi.org/10.1111/sjop.12216.

Soukara, S., Bull, R., Vrij, A., Turner, M., & Cherryman, J. (2009). What really happens in police interviews of suspects? Tactics and confessions. *Psychology, Crime & Law, 15*(6), 493–506. https://doi.org/10.1080/10683160802201827.

Swanner, J. K., Meissner, C. A., Atkinson, D. J., & Dianiska, R. E. (2016). Developing diagnostic, evidence-based approaches to interrogation. *Journal of Applied Research in Memory and Cognition, 5*(3), 295–301. https://doi.org/10.1016/j.jarmac.2016.07.001.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). New York: John Wiley & Sons Ltd.

Vrij, A. (2016). Baselining as a lie detection method: Baselining. *Applied Cognitive Psychology, 30*(6), 1112–1119. https://doi.org/10.1002/acp.3288.

Vrij, A., Blank, H., & Fisher, R. P. (2018). A re-analysis that supports our main results: A reply to Levine et al. *Legal and Criminological Psychology*, *23*(1), 20–23. https://doi.org/10.1111/lcrp.12121.

Vrij, A., & Fisher, R. P. (2016). Which lie detection tools are ready for use in the criminal justice system? *Journal of Applied Research in Memory and Cognition, 5*(3), 302–307. https://doi.org/10.1016/j.jarmac.2016.06.014.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22*(1), 1–21. https://doi.org/10.1111/lcrp.12088.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*(2), 110–117. https://doi.org/10.1016/j.jarmac.2012.02.004.

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Lying about flying: The first experiment to detect false intent. *Psychology, Crime & Law, 17*(7), 611–620. https://doi.org/10.1080/10683160903418213.

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11*(3), 89–121. https://doi.org/10.1177/1529100610390861.

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., et al. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior, 33*(2), 159–166.

Vrij, A., Leal, S., Mann, S. A., & Granhag, P. A. (2011). A comparison between lying about intentions and past activities: Verbal cues and detection accuracy. *Applied Cognitive Psychology, 25*(2), 212–218.

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32*(3), 253–265. https://doi.org/10.1007/s10979-007-9103-y.

Warmelink, L., Vrij, A., Mann, S., & Granhag, P. A. (2013a). Spatial and temporal details in intentions: A cue to detecting deception. *Applied Cognitive Psychology, 27*(1), 101–106. https://doi.org/10.1002/acp.2878.

Warmelink, L., Vrij, A., Mann, S., & Granhag, P. A. (2013b). Spatial and temporal details in intentions: A cue to detecting deception: Spatial and temporal details in lie detection. *Applied Cognitive Psychology, 27*(1), 101–106. https://doi.org/10.1002/acp.2878.

Warmelink, L., Vrij, A., Mann, S., Jundi, S., & Granhag, P. A. (2012). The effect of question expectedness and experience on lying about intentions. *Acta Psychologica, 141*(2), 178–183.

Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D., & Fisher, R. P. (2011). Thermal imaging as a lie detection tool at airports. *Law and Human Behavior, 35*(1), 40–48. https://doi.org/10.1007/s10979-010-9251-3.

Warmelink, L., Vrij, A., Mann, S., Leal, S., & Poletiek, F. H. (2013). The effects of unexpected questions on detecting familiar and unfamiliar lies. *Psychiatry, Psychology and Law, 20*(1), 29–35. https://doi.org/10.1080/13218719.2011.619058.

Wilgoren, J., & Wong, E. (2001). After the attacks: United flight 93; On doomed flight, passengers vowed to perish fighting. *New York Times*. Retrieved from http://www.nytimes.com/2001/09/13/us/after-attacks-united-flight-93-doomed-flight-passengers-vowed-perish-fighting.html.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science, 12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology, 14*, 1–59.

The Many Faces of Trustworthiness: Accuracy and Inaccuracy in Predicting Deception from Facial Appearance

John Paul Wilson and Nicholas O. Rule

People must routinely decide whether to trust another person. We assess whether friends will meet us for lunch at the agreed-upon time, whether business partners will hold up their end of a bargain, whether fellow drivers will obey traffic signals, and whether spouses will remain faithful. Many of these decisions are trivial and even somewhat automatic. For example, we tend to assume that other drivers will obey traffic signals because doing so is in their self-interest as well our own. Other trust decisions are less straightforward, though. In many situations, one party may attempt to deceive the other. Thus, it becomes important for the perceiver (who is the target of the potential deception attempt) to determine whether to trust the actor (who may do some deceiving). In this chapter, we discuss one critically important determinant of whether we decide to trust other people or not: their facial appearance. In doing so, we will specifically focus on the accuracy of these judgments: Can people accurately detect deception and deceptive tendencies in others based on first impressions of their facial appearance?

Before we address the question of accuracy, it is important to discuss whether and how people make trustworthiness judgments of others. An expanding

J. P. Wilson (✉)
Department of Psychology, Montclair State University, Montclair, NJ, USA
e-mail: wilsonjoh@montclair.edu

N. O. Rule
Department of Psychology, University of Toronto, Toronto, ON, Canada
e-mail: rule@psych.utoronto.ca

body of literature over the past decade has shown that people rapidly infer others' trustworthiness with a great deal of consensus (Krumhuber et al., 2007; Willis & Todorov, 2006; Zebrowitz, Voinescu, & Collins, 1996). That is, people tend to agree with one another about who looks trustworthy and who looks untrustworthy, and come to this agreement quickly based on very little information. In fact, some have argued that trustworthiness is one of two basic dimensions that people use to perceive each other (Oosterhof & Todorov, 2008). People evaluate neutral expression faces according to whether they essentially appear to be "good" or "bad" in as little as 38 milliseconds (Bar, Neta, & Linz, 2006), and such good–bad ratings tend to map closely to the slightly more complex impression of whether a person can be trusted (Todorov, Said, Engell, & Oosterhof, 2008).

In considering whether impressions of trustworthiness are accurate, it is also important to address whether they matter. That is, we know that people make snap judgments about trustworthiness, but just how powerful are these impressions in driving social outcomes? Perhaps alarmingly, quite so. People's judgments of the trustworthiness of a face actually predict a wide range of outcomes. Put rather simply, we tend to trust those who have a trustworthy face. For instance, research participants in laboratory-based economic games invest more money with partners who have trustworthy-looking faces (Rezlescu, Duchaine, Olivola, & Chater, 2012; Van't Wout & Sanfey, 2008). In complement, people may enjoy greater success receiving loans if they have a trustworthy face (Duarte, Siegel, & Young, 2012). Trustworthy-looking managers also achieve higher positions in corporations (Linke, Saribay, & Kleisner, 2016). Evidently, then, people tasked with deciding whether a person can be trusted with money and corporate responsibilities seem to rely on facial appearance to make such decisions. This tendency to trust those who look trustworthy is so powerful that even children as young as five years old are less willing to give money to untrustworthy-looking partners in games (Ewing, Caulfield, Read, & Rhodes, 2015).

Trustworthiness decisions also influence the criminal justice system. Judges and juries that render verdicts and levy punishment have to consider how much they can trust defendants. Lawmakers' goals for impartiality notwithstanding, research has shown that a defendant's facial appearance matters in court. For example, participants in recent studies who reviewed a hypothetical criminal case required less evidence to convict a person who looked untrustworthy (Korva, Porter, O'Connor, Shaw, & ten Brinke, 2013; Porter, ten Brinke, & Gustaw, 2010). As such, judgments about guilt and innocence may partly hinge on whether a defendant looks like someone who can be trusted. More disturbing, we recently found that even capital sentencing decisions may be subject to a similar facial bias. In two papers (Wilson & Rule, 2015, 2016), we found that convicted murderers were more likely to have been sentenced to death (vs. life without parole) if they looked less trustworthy. The appearance of trustworthiness thus seems to impact the most consequential legal decision made by the state: whether to execute someone for a crime.

People therefore use appearance to determine both whether they are being deceived and how to punish others. Although the connection between the trustworthiness of a person's appearance and how likely that person is to be trusted seems relatively straightforward, a more complicated issue arises regarding the accuracy of such appearance-based inferences. A fair amount of research has been devoted to this question and we thus spend the rest of this chapter discussing those findings with the aim of elucidating whether people can assess how deceptive others are from their appearance across multiple behavioral domains.

We note from the outset that we will not spend a great deal of time discussing lie detection. A large literature has documented the extent to which people are capable of detecting lies, including multiple meta-analytic summaries (Bond & DePaulo, 2006; Hartwig & Bond, 2011). This literature has demonstrated that people can detect the active commission of lies slightly better than chance. The current review, however, focuses on whether deceptive tendencies can be detected from one's appearance absent any accompanying behavior.

WHY WOULD FIRST IMPRESSIONS OF DECEPTIVE TENDENCIES BE ACCURATE?

Before discussing the empirical findings, it is helpful to consider why we might expect first impressions of deceptive tendencies to be accurate. First, a large literature has examined the accuracy of many types of social judgments (for a review, see Alaei & Rule, 2016). The ecological theory of social perception argues that impressions of people serve adaptive functions; thus, perceivers accordingly make accurate judgments from minimal social cues (Zebrowitz & Montepare, 2006). Beyond nearly perfect accuracy in perceiving social categories such as age, race, and sex (e.g., Macrae & Martin, 2007), people also tend to show a high degree of accuracy in detecting a number of social traits. For example, people demonstrate accuracy in judging others' kinship status (DeBruine et al., 2009; Kaminski, Dridi, Graff, & Gentaz, 2009), intelligence (Zebrowitz, Hall, Murphy, & Rhodes, 2002), health status (Kalick, Zebrowitz, Langlois, & Johnson, 1998; Tskhay, Wilson, & Rule, 2016), extraversion (Borkenau, Brecke, Möttig, & Paelecke, 2009; Borkenau & Liebler, 1992), socioeconomic status (Bjornsdottir & Rule, 2017), sexual orientation (Tskhay & Rule, 2013), religious identification (Allport & Kramer, 1946; Rule, Garrett, & Ambady, 2010), and political ideology (Samochowiec, Wänke, & Fiedler, 2010; Wilson & Rule, 2014). In fact, accuracy in the detection of relevant social information even extends beyond social categorization and trait information to professional success. For instance, Rule and Ambady (2008a) found that initial impressions of the leadership ability of chief executive officers correlated with the profitability of their companies (see also Rule & Ambady, 2011).

From a social ecological perspective, then, it might also be sensible to imagine that deception is detectable from minimal information. Certainly, the ability to detect who is trustworthy or not is adaptive. But so is the ability to successfully deceive. In this way, evolutionary adaptation may be likened to an arms race, with both deceivers and targets of deception adding abilities to their arsenal (Dawkins & Krebs, 1979). Perhaps it is therefore not so straightforward to assume that deception, and dispositions for deception, would be accurately detectable at minimal acquaintance.

The Multiple Facets of Detecting Trustworthiness from the Face

For the most part, we focus on the detection of deceptive tendencies from static facial cues in this review. When we write of "deceptive tendencies," we refer to some presumed disposition toward behaving deceptively. We specify this to differentiate from the active commission of deception. Sometimes these tendencies are inferred from the measurement of a single behavior (e.g., did the participant cheat in a laboratory task while being secretly observed?). Other times these tendencies will be based on a set of behaviors measured or reported overtime (e.g., self-reported infidelity, penalty minutes over a hockey season). Sometimes the behaviors we describe in this review will directly relate to deception, whereas others are better conceptualized as negative actions related to general trustworthiness that are not necessarily deceptive.

Critically, we argue that the detection of trustworthiness and deceptive tendencies is multi-faceted. Although it would be convenient to assume that the detection of cheating tendencies is identical to the detection of criminality or physical aggression, we argue that there are likely multiple domains of trustworthiness, and that some of these may involve behaviors that are more detectable than others. We believe that it is generally problematic to think of trustworthiness as a unitary construct, partly because trustworthiness, broadly defined, can include most behaviors generally considered good or helpful, and untrustworthiness can include most behaviors generally considered bad or harmful. This general framework could thus include a number of behaviors: aggression (Carré, McCormick, & Mondloch, 2009; Carré, Murphy, & Hariri, 2013; Stillman, Maner, & Baumeister, 2010), criminality (Porter, England, Juodis, ten Brinke, & Wilson, 2008; Rule, Krendl, Ivcevic, & Ambady, 2013), lie detection (Bond & DePaulo, 2006; Hartwig & Bond, 2011), academic cheating (Davis & Ludvigson, 1995; Geniole, Keyes, Carré, & McCormick, 2014; Rule et al., 2013), financial malfeasance (Rule et al., 2013), sexual infidelity (Rhodes, Morley, & Simmons, 2013), and self-interest in economic games (Stirrat & Perrett, 2010; Verplaetse, Vanneste, & Braeckman, 2007).

Although each of these domains arguably involves some facet of trustworthiness, an assessment of the literature surrounding each one suggests that

they are not equally detectable from first impressions. As we have recently argued (Wilson & Rule, 2017), it is likely that the literature contains disparate findings because of the multiple behavioral domains encapsulated by trustworthiness as an umbrella term. We briefly review the domains presently addressed in the literature below.

Cheating and Honesty

Some of the earlier social psychological research that tackled the question of whether trustworthiness and deceptive tendencies can be read from the face did so through the lens of honesty. One striking example comes from Zebrowitz et al. (1996), who investigated the relationship between facial appearance and honesty across the life span. Using data from the Intergenerational Studies of Development and Aging, they recruited judges to code the faces of the study's participants for attractiveness, babyfacedness, facial expression, and perceived honesty. Critically, the individuals had been assessed for actual honesty by professional clinicians at multiple time points across the study's roughly 50-year span. Zebrowitz et al. found an interesting relationship between actual and perceived honesty. Among men, they observed a self-fulfilling prophecy, such that men who looked more honest at younger ages became more honest later. For women, though, they observed what they called an "artifice effect." Women who were actually less honest early in life came to appear more honest later in life, perhaps suggesting that they altered their appearance to more effectively take advantage of their dishonest tendencies.

Other attempts to plumb the association between honesty and facial appearance have occurred on a smaller scale. Bond, Berry, and Omar (1994) found that people perceived as dishonest-looking were actually more likely to volunteer to participate in research requiring them to act deceptively than were people who looked more honest. This suggests that willingness to deceive is perhaps apparent from one's appearance. More recently, however, Rule et al. (2013) recruited participants to first sit for a photograph and then complete an ostensibly unrelated problem-solving task. Participants were asked to complete as many math and language test questions as they could within a five-minute period. They were given a timer and asked to stop working when the timer expired. Unaware that they were being discreetly observed, participants could cheat by continuing to work after the timer expired. A separate group of judges rated the trustworthiness of each participant's face from the photographs taken at the beginning of the session. Here, ratings of trustworthiness from photographs did not relate to cheating behavior—neither whether they cheated nor how long they cheated. As such, the association between honesty and one's appearance is far from straightforward in this subdomain, with systematic inaccuracy (Zebrowitz et al., 1996) perhaps just as likely as systematic accuracy.

Infidelity

Related work has looked at another very common version of everyday trust-worthiness: sexual infidelity. Only a few studies have examined whether sexual infidelity tendencies are detectable from the face, but there are some indications that they are. Rhodes et al. (2013), for example, found that women could judge men's unfaithfulness from facial photographs better than chance. Men could not accurately judge women's unfaithfulness, however. These researchers also found divergences between perceptions of unfaithfulness and perceptions of trustworthiness. Despite an apparently strong link between these two constructs, only perceptions of unfaithfulness actually related to self-reports of infidelity. Follow-up work found that men may have a limited ability to detect unfaithfulness in the faces of women under certain conditions (Leivers, Simmons, & Rhodes, 2015).

Behavior in Economic Games

Psychologists have become increasingly interested in people's behavior in economic decision-making games. Such behavior is particularly appropriate for the present topic because the decisions often take the form of face-valid trust behavior. For example, in the Trust Game (Berg, Dickhaut, & McCabe, 1995), one participant in a pair must decide how much money to entrust in the other participant, with the assumption that trustworthy participants will return a fair share of the money after it has been multiplied according to the rules of the game. Other economic games measure similar behaviors. For example, some work has investigated whether facial appearance can predict behavior in the prisoner's dilemma game, in which participants choose whether to cooperate or defect in a dyadic scenario. In one study, Verplaetse et al. (2007) determined that participants could identify noncooperative participants in such a game more accurately than chance. However, they could only do so when viewing photographs taken at the moment that the cooperation decision was made. Such a result suggests that there is likely nothing in a person's static facial appearance that signals their trustworthiness or cooperativeness. Rather, perhaps people express subtle signals that reveal their intentions as they make the decision. More recent research using Verplaetse et al.'s stimuli found below-chance accuracy for detecting noncooperation (Shoda & McConnell, 2013), suggesting that any accuracy for deception detection with this particular sample is inconsistent.

Other research has found that such decisions may be predicted by the structural information contained in a person's face. In particular, researchers have looked at facial width-to-height ratio (fWHR), a measure of how wide a person's face is relative to its height that positively relates to testosterone levels in men (Lefevre, Lewis, Perrett, & Penke, 2013). Stirrat and Perrett (2010) found that fWHR among men predicted their tendency to take advantage of others in a trust game. Furthermore, people were less likely to trust male partners with wider faces. Related work found that fWHR

positively correlated with the tendency to exploit partners in a negotiation (Haselhuhn & Wong, 2011).

Although this fWHR research suggests that the face does contain signals that broadcast its owner's deceptive tendencies, other findings have been more nuanced and have suggested that perceivers do not always glean valuable information from the face. Some of this work has shown that explicit judgments of trustworthiness are not associated with cooperative behavior, and that maybe only implicit trustworthiness judgments are accurate. In one set of studies, Bonnefon, Hopfensitz, and De Neys (2013) photographed people who played the role of trustee in a trust game and, as such, had made a decision about whether or not to exploit the trust of their partners. They showed the photographs to new participants and asked them to decide whether to trust these trustees by investing money with them, in the hope that these investors would send back some of the money from the new greater total. Participants entrusted less money to trustees who, unbeknownst to them, had chosen to exploit their previous partners' trust, showing that they had some intuition about who was trustworthy and who was not. Other participants who were simply asked to rate each trustee's face for trustworthiness did not show accurate judgments; however, explicit trustworthiness judgments were uncorrelated with the total amount returned by trustees. Bonnefon et al. also found that the amount of facial information presented to perceivers impacted accuracy in a surprising way: People who viewed photographs containing *more* information, including hairstyle and clothing, were *less* accurate in judging actual trustworthiness than those who only saw core facial features (i.e., the internal features of the face such as eyes, ears, and mouth).

Even more recent work has shown that independent judges' ratings of others' trustworthiness in a mixed-motive game may predict actual behavior. In this work, Slepian and Ames (2016) recruited M.B.A. students to play a two-person game in which they chose to either lie or tell the truth, and to then spend a few minutes persuading their partner that they were being truthful. Lying successfully was incentivized in the game. Judges who subsequently rated photographs of the participants' faces accurately detected their trustworthiness slightly better than chance. Moreover, targets' expectations of how much people would trust them mediated the link between judges' ratings and their actual trustworthy behavior. In particular, people who expected to be trusted tended to behave in a more trustworthy manner, and people who expected not to be trusted were more deceptive. These findings offer evidence that, in some situations, people may internalize others' expectations about their trustworthiness and behave accordingly.

Physical Aggression

Quite a bit of research in recent years has explored whether and how facial structure relates to physically aggressive behavior, again focusing on fWHR. In one study, Carré and McCormick (2008) found that fWHR positively predicted both self-reported trait dominance and the tendency to aggress against

fellow participants in a laboratory-based measure of aggression. In the same paper, these researchers found that fWHR also positively related to the average number of penalty minutes a group of university hockey players incurred per game across a full season. Perceivers may employ these valid aggression cues as well. For example, Carré et al. (2009) found that estimates of aggression strongly correlated with targets' fWHR and, in turn, positively correlated with actual aggression.

Although not all studies have found that fWHR relates to the propensity to engage in aggressive behavior (Deaner, Goetz, Shattuck, & Schnotala, 2012; Gómez-Valdés et al., 2013; Özener, 2012), the weight of existing evidence appears to favor the perspective that fWHR is a valid aggression cue. This link has been confirmed by studies showing that fWHR positively predicts fighting ability among mixed martial arts athletes (Zilioli et al., 2015) and penalties incurred by soccer players (Welker, Goetz, Galicia, Liphardt, & Carré, 2015). Overall, meta-analysis of the extant research confirms that aggressive tendencies can be extracted from facial information with some accuracy (Geniole, Denson, Dixson, Carré, & McCormick, 2015).

Though not all studies looking at physical aggression have involved measurements of facial structure, they may nevertheless confirm that physical aggression can be gleaned from the face. In one study, participants viewing mug shots of violent and nonviolent criminals could differentiate between the two groups more accurately than chance (Stillman et al., 2010). As we noted above, the links between testosterone and facial appearance may help to explain why some types of untrustworthy behaviors (e.g., aggression and violence) are detectable from one's appearance.

Criminality

Despite the links between appearance and aggression, whether a more general tendency to engage in criminal behavior can be detected from the face remains a separate question. Some research has suggested that people can differentiate criminals from noncriminals. For example, Porter et al. (2008) showed participants photographs of people who had appeared on a list of *America's Most Wanted* criminals and those who had received the Nobel Peace Prize. Although participants could accurately differentiate between the two groups, the study is limited by its comparison of drastically different photograph types (i.e., criminal mug shots vs. professional photographs taken from the Nobel website). In a similar study, Rule et al. (2013) collected mug shots of Hollywood celebrities as well as professional photographs of the same individuals. Participants could not distinguish Nobel Prize winners from the celebrities when their professional photographs were used, but could when viewing their mug shots. This suggests that when the source of the photograph is equated, criminality cannot necessarily be read from the face. In separate additional studies published in the same paper, Rule et al. found that judges could not differentiate convicted US war criminals from

decorated military heroes, and also that executives convicted of financial crimes looked no less trustworthy than executives not convicted of crimes. Across multiple types of untrustworthy criminal behavior, then, trustworthiness and deception do not seem to be easily detected from one's appearance.

Trustworthiness as a Trait Versus State

It merits noting that the literature we have covered in this chapter is generally concerned with trait trustworthiness and deceptiveness. This can be differentiated from state trustworthiness (e.g., engaging in individual deception attempts). It is of course true that much research has investigated individual acts of deception and people's ability to detect them, with summaries showing that people can detect deception attempts better than chance (but with nowhere near perfect accuracy). For example, one meta-analysis estimated that people can tell truths from lies about 54% of the time (Bond & DePaulo, 2006). Other analyses have shown that the behaviors that differentiate truth-tellers from liars tend to be quite small and poorly identified by perceivers (Hartwig & Bond, 2011).

Here, we have focused on definitions of trustworthiness and deceptiveness that may be best thought of as a trait. Of course, many of the operationalizations in the reviewed literature are characterized by one specific behavior, perhaps even just at one time point (e.g., cheating on a task in the laboratory, committing a financial or violent crime). There is at least some evidence that those who cheat once will tend to cheat again, however. Davis and Ludvigson (1995) found that nearly all undergraduates who repeatedly committed acts of academic dishonesty reported having cheated in high school. Rule et al. (2013) found that people who cheated in a laboratory task were more likely to claim that they would never cheat, and Zebrowitz et al. (1996) found that real honesty tended to be relatively stable overtime.

The Importance of Defining Trustworthiness

Although it is likely clear to the reader that deceptive tendencies and general trustworthiness are not reliably detectable from one's appearance, it may not be quite so clear why we consider it important to think more carefully about the multiple domains of trustworthiness. Consider, though, the way in which questions about appearance-based judgments are often posed. Scholars and laypeople often ask some variant of the question "Can you read a book by its cover?" when posing questions about the informational value of one's appearance. There are, of course, some long-standing beliefs about the link between appearance and behavior, dating back to early writings on physiognomy by Lavater (1800). Some contemporary scholars would argue that, by and large, most social attributions from faces are inaccurate (e.g., Todorov, Olivola, Dotsch, & Mende-Siedlicki, 2015). Others have argued that there are "kernels of truth" in social inferences based on

facial appearance (Bonnefon, Hopfensitz, & De Neys, 2015). This latter perspective relies on some of the same evidence that we have reviewed here regarding economic game behavior (Bonnefon et al., 2013; see also Little, Jones, DeBruine, & Dunbar, 2013).

The perspective that there is some accuracy in social attributions from faces is supported by a large and growing literature. Whether people are judging others on the basis of personality (Borkenau & Liebler, 1992), health (Tskhay et al., 2016), socioeconomic status (Bjornsdottir & Rule, 2017), or social identities like sexual orientation (Rule & Ambady, 2008b) and religious affiliation (Allport & Kramer, 1946), people do tend to have the ability to judge *something* about the book based on its cover. When it comes to trustworthiness and the tendency to deceive, however, there is no simple answer. We argue that scholars benefit from asking the question in a more nuanced and focused fashion, emphasizing not *whether* we can judge someone's trustworthiness from their appearance, but under what circumstances and to what extent.

By and large, people seem to lack the ability to glean accurate information about trustworthiness and deceptiveness from the face. Cases in which there is some accuracy seem to come mostly from domains involving aggression. This suggests that whereas behavioral tendencies that have a strongly physiological basis may be accurately detected, those that involve purer forms of deception may not. However, this supposition has not yet been strongly confirmed by existing research. Moving forward, researchers should conduct studies with these potential distinctions in mind in order to come to a better specified taxonomy of trustworthiness and deception. "Trustworthiness" may be quite effective at capturing a fundamental domain of person perception, but it is lacking as an umbrella trait for actual behavior.

It is important to know how accurate perceptions of trustworthiness and deception are because, in many ways, your face is your fate. From political decision-making to leadership selection to even life-and-death capital punishment decisions, people's facial appearance plays a considerable role in how they are treated by others. Some appearance-based inferences have some validity but, as we have reviewed, others do not. When one attempts to decide whether he or she is being deceived, it is important to know whether first impressions can be trusted. Researchers and laypeople alike would benefit from knowing when these impressions are likely to be invalid and when they may, in fact, hold a small kernel of truth. A better taxonomy of trustworthiness will move us in the right direction.

References

Alaei, R., & Rule, N. O. (2016). Accuracy of perceiving social attributes. In J. A. Hall, M. Schmid Mast, & T. V. West (Eds.), *The social psychology of perceiving others accurately* (pp. 125–142). Cambridge: Cambridge University Press.

Allport, G. W., & Kramer, B. M. (1946). Some roots of prejudice. *The Journal of Psychology, 22,* 9–39.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6,* 269–278.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10,* 122–142.

Bjornsdottir, R. T., & Rule, N. O. (2017). The visibility of social class from facial cues. *Journal of Personality and Social Psychology, 113,* 530–546.

Bond, C. F., Jr., Berry, D. S., & Omar, A. (1994). The kernel of truth in judgments of deceptiveness. *Basic and Applied Social Psychology, 15,* 523–534.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234.

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General, 142,* 143–150.

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2015). Face-ism and kernels of truth in facial inferences. *Trends in Cognitive Sciences, 19,* 421–422.

Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality, 43,* 703–706.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62,* 645–657.

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society of London B: Biological Sciences, 275,* 2651–2656.

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science, 20,* 1194–1198.

Carré, J. M., Murphy, K. R., & Hariri, A. R. (2013). What lies beneath the face of aggression? *Social Cognitive and Affective Neuroscience, 8,* 224–229.

Davis, S. F., & Ludvigson, H. W. (1995). Additional data on academic dishonesty and a proposal for remediation. *Teaching of Psychology, 22,* 119–121.

Dawkins, R., & Krebs, J. R. (1979). Arms races between and within species. *Proceedings of the Royal Society of London B: Biological Sciences, 205,* 489–511.

Deaner, R. O., Goetz, S. M., Shattuck, K., & Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *Journal of Research in Personality, 46,* 235–238.

DeBruine, L. M., Smith, F. G., Jones, B. C., Roberts, S. C., Petrie, M., & Spector, T. D. (2009). Kin recognition signals in adult faces. *Vision Research, 49,* 38–43.

Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies, 25,* 2455–2484.

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science, 18,* 327–334.

Geniole, S. N., Denson, T. F., Dixson, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PLoS ONE, 10*(7), e0132726.

Geniole, S. N., Keyes, A. E., Carré, J. M., & McCormick, C. M. (2014). Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Personality and Individual Differences, 57,* 59–64.

Gómez-Valdés, J., Hünemeier, T., Quinto-Sánchez, M., Paschetta, C., de Azevedo, S., González, M. F., … Bau, C. H. (2013). Lack of support for the association between facial shape and aggression: A reappraisal based on a worldwide population genetics perspective. *PLoS ONE, 8*(1), e52317.

Hartwig, M., & Bond, C. F., Jr. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137,* 643–659.

Haselhuhn, M. P., & Wong, E. M. (2011). Bad to the bone: Facial structure predicts unethical behaviour. *Proceedings of the Royal Society of London B: Biological Sciences.* https://doi.org/10.1098/rspb20111193.

Kalick, S. M., Zebrowitz, L. A., Langlois, J. H., & Johnson, R. M. (1998). Does human facial attractiveness honestly advertise health? Longitudinal data on an evolutionary question. *Psychological Science, 9,* 8–13.

Kaminski, G., Dridi, S., Graff, C., & Gentaz, E. (2009). Human ability to detect kinship in strangers' faces: Effects of the degree of relatedness. *Proceedings of the Royal Society of London B: Biological Sciences, 276,* 3193–3200.

Korva, N., Porter, S., O'Connor, B. P., Shaw, J., & ten Brinke, L. (2013). Dangerous decisions: Influence of juror attitudes and defendant appearance on legal decision-making. *Psychiatry, Psychology and Law, 20,* 384–398.

Krumhuber, E., Manstead, A. S., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion, 7*(4), 730–735.

Lavater, J. C. (1800). *Essays on physiognomy: For the promotion of the knowledge and the love of mankind; written in the German language by JC Lavater, abridged from Mr. Holcrofts translation* (Printed for GGJ & J. Robinson).

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior, 34,* 273–279.

Leivers, S., Simmons, L. W., & Rhodes, G. (2015). Men's sexual faithfulness judgments may contain a kernel of truth. *PLoS ONE, 10,* e0134007.

Linke, L., Saribay, S. A., & Kleisner, K. (2016). Perceived trustworthiness is associated with position in a corporate hierarchy. *Personality and Individual Differences, 99,* 22–27.

Little, A. C., Jones, B. C., DeBruine, L. M., & Dunbar, R. I. (2013). Accuracy in discrimination of self-reported cooperators using static facial information. *Personality and Individual Differences, 54,* 507–512.

Macrae, C. N., & Martin, D. (2007). A boy primed Sue: Feature-based processing and person construal. *European Journal of Social Psychology, 37,* 793–805.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105,* 11087–11092.

Özener, B. (2012). Facial width-to-height ratio in a Turkish population is not sexually dimorphic and is unrelated to aggressive behavior. *Evolution and Human Behavior, 33,* 169–173.

Porter, S., England, L., Juodis, M., ten Brinke, L., & Wilson, K. (2008). Is the face a window to the soul? Investigation of the accuracy of intuitive judgments of the trustworthiness of human faces. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 40,* 171−177.

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law, 16,* 477–491.

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PLoS ONE, 7*(3), e34293.

Rhodes, G., Morley, G., & Simmons, L. W. (2013). Women can judge sexual unfaithfulness from unfamiliar men's faces. *Biology Letters, 9,* 20120908.

Rule, N. O., & Ambady, N. (2008a). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science, 19,* 109–111.

Rule, N. O., & Ambady, N. (2008b). Brief exposures: Male sexual orientation is accurately perceived at 50-ms. *Journal of Experimental Social Psychology, 44,* 1100–1105.

Rule, N. O., & Ambady, N. (2011). Face and fortune: Inferences of personality from Managing Partners' faces predict their law firms' financial success. *The Leadership Quarterly, 22,* 690–696.

Rule, N. O., Garrett, J. V., & Ambady, N. (2010). Places and faces: Geographic environment influences the ingroup memory advantage. *Journal of Personality and Social Psychology, 98,* 343–355.

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology, 104,* 409–426.

Samochowiec, J., Wänke, M., & Fiedler, K. (2010). Political ideology at face value. *Social Psychological and Personality Science, 1,* 206–213.

Shoda, T. M., & McConnell, A. R. (2013). Interpersonal sensitivity and self-knowledge: Those chronic for trustworthiness are more accurate at detecting it in others. *Journal of Experimental Social Psychology, 49,* 440–443.

Slepian, M. L., & Ames, D. R. (2016). Internalized impressions: The link between apparent facial trustworthiness and deceptive behavior is mediated by targets' expectations of how they will be judged. *Psychological Science, 27,* 282–288.

Stillman, T. F., Maner, J. K., & Baumeister, R. F. (2010). A thin slice of violence: Distinguishing violent from nonviolent sex offenders at a glance. *Evolution and Human Behavior, 31,* 298–303.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust male facial width and trustworthiness. *Psychological Science, 21,* 349–354.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66,* 519–545.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12,* 455–460.

Tskhay, K. O., & Rule, N. O. (2013). Accuracy in categorizing perceptually ambiguous groups: A review and meta-analysis. *Personality and Social Psychology Review, 17,* 72–86.

Tskhay, K. O., Wilson, J. P., & Rule, N. O. (2016). People use psychological cues to detect physical disease from faces. *Personality and Social Psychology Bulletin, 42,* 1309–1320.

Van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition, 108*, 796–803.

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: The sequel—A kernel of truth in predictive cheating detection. *Evolution and Human Behavior, 28,* 260–271.

Welker, K. M., Goetz, S. M., Galicia, S., Liphardt, J., & Carré, J. M. (2015). An examination of the associations between facial structure, aggressive behavior, and performance in the 2010 World Cup Association Football Players. *Adaptive Human Behavior and Physiology, 1,* 17–29.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17,* 592–598.

Wilson, J. P., & Rule, N. O. (2014). Perceptions of others' political affiliation are moderated by individual perceivers' own political attitudes. *PLoS ONE, 9*(4), e95431.

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science, 26,* 1325–1331.

Wilson, J. P., & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: The role of facial trustworthiness. *Social Psychological and Personality Science, 7,* 331–338.

Wilson, J. P., & Rule, N. O. (2017). Advances in understanding the detectability of trustworthiness from the face: Toward a taxonomy of a multifaceted construct. *Current Directions in Psychological Science, 26,* 396–400.

Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin, 28,* 238–249.

Zebrowitz, L. A., & Montepare, J. M. (2006). The ecological approach to person perception: Evolutionary roots and contemporary offshoots. In M. Schaller, J. A. Simpson, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 81–113). New York: Psychology Press.

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-Eyed" and "Crooked-Faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin, 22,* 1258–1269.

Zilioli, S., Sell, A. N., Stirrat, M., Jagore, J., Vickerman, W., & Watson, N. V. (2015). Face of a fighter: Bizygomatic width as a cue of formidability. *Aggressive Behavior, 41,* 322–330.

Contexts of Deceptive Communication: The Self and Identity

Self-Assessed Lie- and Truth-Telling Abilities: Demographic, Personality, and Behavioral Correlates

Eitan Elaad

People use their ability to convince other people of their truthful communication to safeguard social relationships, to ensure financial advantages, and even to gain trust in law enforcement situations. At the same time, people lie from time to time (Ariely, 2012) and the perceived ability to be persuasive when lying inspires the senders' success in these situations. In this context, the concept of self-efficacy is central. Bandura (1977, 1992) defined self-efficacy as people's belief in their ability to accomplish goals in given situations. Bandura described these beliefs as determinants of how people think, behave, and feel. Following the self-efficacy model, people's confidence in their ability to convince when they are truthful and persuasive when they are lying may be important in contexts of interpersonal communication. However, research on various aspects of the self-assessed lie- and truth-telling abilities is in its infancy. The following is a review of the related literature and offers direction for future research.

E. Elaad (✉)
Department of Behavioral Sciences, Ariel University, Ariel, Israel
e-mail: elade@ariel.ac.il

High Assessments of the Ability to Convince When Telling the Truth and Low Assessments of the Ability to Lie Persuasively

Early accounts suggest that individuals tend to rate their own lie-telling ability relatively low. Elaad (2003) reported that police interrogators and police personnel, using a nine-point scale, gave low ratings to their own lie-telling ability (Mean = 4.45, SD = 2.05), below the middle point (5) "as good as others". A series of studies, summarized in Table 23.1, compare participants' self-assessed lie- and truth-telling abilities. Participants in these studies were asked: "In comparison with other people, how would you assess your own ability to tell the truth convincingly?" and "In comparison with other people, how would you assess your own ability to tell lies persuasively?" Answers were given on a scale ranging from 0 (*much worse than others*) to 100 (*much better than others*), with 50 (*as good as others*) serving as the middle point.

To summarize the differences between the self-assessed truth-telling ability and the self-assessed lie-telling ability, a mini meta-analysis was performed on the results presented in Table 23.1. Mini meta-analyses have been used previously (e.g., Lamarch & Murray, 2014; Williams & DeSteno, 2008) and have recently gained support (e.g., Goh, Hall, & Rosenthal, 2016; Maner, 2014). Goh et al. (2016) contended that even with a small number of studies (actually two studies are sufficient), a meta-analytic procedure allows us to succinctly summarize results across studies and clarify emerging insights.

Table 23.1 indicates that people tend to self-assess their ability to convincingly tell the truth higher than their ability to lie convincingly. This is clearly indicated by the large weighted mean d_{RM}. In support to this conclusion, Table 23.1 further shows that 15 of the 16 lie- and truth-telling ability comparisons are significantly different.

Another important question is whether individuals' self-assessed lie-telling ability is underestimated whereas individuals' truth-telling ability is overestimated. To clarify the assumed biases, a 95% confidence interval (CI) was computed for each assessed ability within each condition. CIs provide useful information about differences between the self-assessed ability and the middle point: "as good as others". To assure that the mean ability rating is not just a sampling error, the CI was based on standard error units. Thus, if the lower bound of the CI is larger than the middle point (.50), confirmation for the overestimation of self-assessed ability is provided. Similarly, if the upper bound of the CI is smaller than the middle point, it may be assumed that the self-assessed ability is underestimated. To estimate if the bias is systematic across all conditions, a correction for sampling error was applied (Hunter & Schmidt, 1990), and the remaining variance, after the correction, was used to compute the 95% CI for the weighted mean statistic. In cases where the remaining variance (correct variance) is too small or equals 0, the mean can be assumed to be the correct value and no CI was computed.

Table 23.1 Percent means (and SDs), of self-assessed abilities to tell lies and truths

	Truth-telling			Lie-telling						
	Mean	SD	95% CI	Mean	SD	95% CI	N	t (paired)	d$_{RM}$	r
Elaad (2006)										
Secular students	.73	(.18)	.675–.778	.50	(.23)	.432–.568	48	6.8*	0.99	.40
Secular kibbutz	.62	(.16)	.576–.668	.53	(.28)	.449–.607	50	1.9	0.26	−.28
Religious 1	.71	(.17)	.654–.761	.38	(.28)	.294–.471	40	6.9*	1.09	.18
Religious 2	.66	(.16)	.614–.705	.33	(.24)	.265–.400	52	8.8*	1.22	.18
Elaad (2009)										
Laypersons	.73	(.19)	.656–.797	.48	(.27)	.374–.576	30	4.9*	0.89	.29
Prisoners	.82	(.12)	.772–.864	.44	(.22)	.354–.519	30	9.1*	1.65	.20
Interrogators	.76	(.13)	.711–.811	.63	(.19)	.556–.702	28	4.5*	0.86	.58
Elaad et al. (2012)										
Adolescents	.68	(.18)	.648–.714	.49	(.28)	.443–.543	121	7.1*	0.64	.24
Elaad (2015a)										
Students	.80	(.21)	.749–.841	.37	(.30)	.297–.433	80	10.9*	0.81	.09
Elaad (2015b)										
Prosecutors	.76	(.17)	.702–.823	.42	(.25)	.326–.509	32	8.7*	1.53	.49
Lay-people	.69	(.19)	.618–.756	.45	(.25)	.355–.535	32	5.1*	0.91	.30
Elaad and Reizer (2015)										
Students	.66	(.18)	.633–.686	.46	(.27)	.418–.499	174	9.6*	0.72	.28
Elaad and Sommerfeld (2016)										
Students	.68	(.17)	.647–.713	.48	(.20)	.435–.515	100	9.2*	0.92	.29
Yaacov (2017)										
Community	.68	(.18)	.656–.706	.38	(.25)	.346–.416	192	14.6*	1.05	.12
Elaad (2018)										
Secular	.79	(.15)	.752–820	.67	(.21)	.624–.716	80	5.5*	0.57	.38
Religious	.68	(.16)	.640–.713	.39	(.17)	.385–.400	80	12.2*	1.36	.21
Weighted *Means*	.71			.46					0.91	.22

*$p < .001$, two tailed
This is a repeated measures design which measures differences within a person. Therefore, repeated measure effect size (d_{RM}) was used (see, Morris & DeShon, 2002) as follows:

$$d_{RM} = \frac{M\,t - M\,l}{SD\,d}$$

Where $M\,t$ is the mean truth telling assessment and $M\,l$ is the mean lie-telling assessment
95% CIs were defined in standard error units

Inspection of Table 23.1 suggests that individuals' truth-telling ability is overestimated. The lower bounds of all 16 CIs are clearly above the middle point. After applying the correction for sampling error, it was found that the correct variance, the difference between the mean variance and the sample error variance, equals 0. The weighted mean across conditions was assumed to be the correct value and no CI was added. Table 23.1 shows that the weighted truth-telling mean is well above .50 and therefore overestimated. The results for the lie-telling ability assessments are not consistent, and the upper bound of the CI computed for the weighted mean is above .50. This leads to the conclusion that the lie-telling ability assessment is not biased.

However, very different populations were used in the different conditions (e.g., religious people, secular, students, prisoners, and prosecutors) which raises the question whether the overestimation of the lie-telling ability by some groups (e.g., interrogators) overshadow the underestimated self-assessed lie-telling ability of other groups. Furthermore, the upper bounds of 6 CIs are below .50, indicating that the lie-telling ability is underestimated. Half of these CIs were computed for religious groups. A search for moderators can be justified.

Participants who are overconfident in their ability to convince receivers of their truthful communication appear to follow the line of reasoning that truth-telling is a simple matter of "telling it like it is" (Buller & Burgoon, 1996; Miller & Stiff, 1993). This confidence also fits general human assumption that most communications are truthful (i.e., truth bias) and there is no reason for others to doubt our own truthful messages.

The truth-telling bias can also be explained by the *illusion of transparency* (Gilovich, Savitsky, & Medvec, 1998), which suggests that in communications, senders are anchored to their own internal experience. Although senders realize that recipients are not exposed to the same information as they are, their adjustment is insufficient (Tversky & Kahneman, 1974), yet they nonetheless tend to believe that receivers can discern their internal states and trust them when they tell the truth. Finally, the high self-assessed ability to be convincing when telling the truth supports the human desire to sustain one's positive self-image (Kaplar & Gordon, 2004). Individuals who rate their truth-telling ability as highly skilled serve this self-image.

The illusion of transparency (Gilovich et al., 1998) may also explain the lower lie-telling ratings that were reported. People mistakenly believe that their lies shine through (Vrij, 2008). Although senders realize that others are unable to perceive their inner-feelings as they do, their adjustment is insufficient, and senders fear that receivers can detect their lies. The desire to sustain a positive self-image also contributes to the relatively low lie-telling ratings, as individuals who are poor at telling lies may believe that they are honest people.

GROUP DIFFERENCES IN LIE- AND TRUTH-TELLING ABILITY ASSESSMENTS

Table 23.1 highlights some possible moderators that may account for group differences in lie- and truth-telling ability assessments, including religiosity, gender, age, and in-service lying experience.

Religiosity

The relation between religiosity and lie- and truth-telling abilities has been studied in a few settings. Elaad's (2006) work, published in Hebrew, included four groups of Israelis, comprised of secular students, secular kibbutz

members, religious individuals (religious 1), and members of a segregated Jewish religious community (religious 2). It was hypothesized that religious people may be less cognitively flexible than their secular counterparts. Cognitive flexibility is the ability to restructure knowledge in multiple ways depending on changing situational demands (Spiro, Feltovich, Jacobson, & Coulson, 1992). In the context of deception, Ariely (2012) used the term cognitive flexibility to describe how people manage to live with two conflicting motivations: On the one hand, they want to benefit from cheating, yet on the other hand, they continue to consider themselves honest people, so they "fudge". Results from Elaad's (2006) study indicated that while all groups overestimated their own truth-telling ability, religious people tended to rate their lie-telling ability significantly lower than secular people.

A recent study (Elaad, 2018) reexamined the effect of religiosity on lie- and truth-telling ability assessments (Table 23.1). Participants were 80 religious and 80 secular people from the local community. Results indicated that religious participants rated their lie-telling ability lower than secular participants. Religious participants also rated their truth-telling ability significantly lower than their secular counterparts.

Further, across the three religious conditions, the weighted mean lie-telling ability assessment was .37. Since all the variance is expected from sampling error, the weighted mean is assumed as the correct value. The corresponding secular conditions produced a weighted mean lie-telling ability assessment of .58 and a 95% CI of .466 to .703, leading to the conclusion that religious participants underestimated their lie-telling ability whereas secular participants did not. As to the truth-telling ability, the three religious groups yielded a weighted mean of .68. It was found that all the variance is expected from sampling error and therefore the computed weighted mean is treated as the correct value. The three corresponding secular groups exhibited a weighted mean of .68 and a respective 95% CI of .662 to .763.

It appears that religiosity moderates lie-telling ability assessments to a point where lie-telling is clearly underestimated. Both religious and secular participants overestimated their truth-telling ability. The moderating effect of religiosity on lie-telling ability assessments is explained by Jewish religious rules that condemn lying. To comply with these rules, religious people underrate their lie-telling ability, allowing them to preserve their honest self-image.

Gender

Gender differences may also play a role in assessing lie-truth communication skills. A meta-analysis of scales from widely used personality inventories from 1940 to 1992 shows that females scored slightly but consistently higher on scales of trust (Feingold, 1994). Such trust reflects a belief in the honesty and positive intentions of individuals, experts, public officials, the media, and others. Females also reported telling fewer lies than males and scored lower

on the Social Adroitness scale, which was designed to identify ambitious persons skilled in subtle, diplomatic persuasion (Kashy & DePaulo, 1996). These findings suggest that females are more sensitive than males to honesty and therefore may evaluate their lie-telling ability lower than males. The more ambitious males may rate their lie-telling ability higher than females.

Gender differences were also examined in Elaad's (2015a) more recent study. In this study, 40 male and 40 female students self-assessed their lie- and truth-telling abilities. Results indicated that female students rated their lie-telling ability slightly lower than male students, but the difference did not reach statistical significance. In additional recent studies, Elaad and Reizer (2015) and Elaad (2018) provided reexamined gender differences. Elaad and Reizer (2015) found a significant difference between the lie-telling assessments of 84 male students (Mean 50.0, SD = 28.0) and 88 female students (Mean 40.9, SD = 25.2) ($t_{(170)} = 2.37$, $p = .019$, $d = .34$). However, the effect size is small and dictates caution. No gender differences were found for truth-telling assessments. Data extracted from Elaad's (2018) study enable a comparison of the lie-telling ability assessments of 79 males from the local community (Mean 57.2, SD = 26.0) and 81 females (Mean 48.9, SD = 20.4). Again, the difference was significant ($t_{(158)} = 2.25$, $p = .025$, $d = .36$) with a small effect size. No significant gender differences were observed for the truth-telling assessments.

All three studies indicate that compared with male participants, females gave somewhat lower assessments to their lie-telling ability, yet the difference is either insignificant or has a small effect size. Furthermore, studies using a male majority of participants (e.g., Elaad, 2009) display similar lie- and truth-telling assessments compared to studies that used a female majority (e.g., Elaad & Sommerfeld, 2016).

Gender biases in the assessments of lie- and truth-telling abilities were examined by computing weighted means and the corresponding 95% CIs across the three studies (Elaad, 2015a, 2018; Elaad & Reizer, 2015). Females underestimated their lie-telling ability (weighted mean = .42, 95% CI = .373 to .472) and overestimated their truth-telling ability (weighted mean = .70, 95% CI = .636 to .756). Males overestimated their truth-telling ability (weighted mean = .73, 95% CI = .645 to .819), but not their lie-telling ability (weighted mean = .51, 95% CI = .459 to .562). Therefore, both males and females are biased toward enhanced truth-telling ability assessments. Only females underestimated their lie-telling ability. Nevertheless, the effect of gender on the self-assessed lie-telling ability is not yet conclusive and requires additional investigation.

Age

Another unresolved question is how the self-assessed truth-telling and lie-telling skills change over individuals' lifespan. A recent survey on deceptive behavior occurring over the lifespan (Debey, Schryver, Logan, Suchotzki, &

Verschuere, 2015) indicated that lying proficiency and frequency decrease in adulthood. Seniors (over age 60) were the worst and least frequent liars of all examined adult groups. Other studies that tested lie-telling skills in older adults (e.g., Ruffman, Murray, Halberstadt, & Vater, 2012) reported that older adults (60–89 years old) were more easily detected when they lied, than young adults (17–26 years old). It is believed that as they become economically and mentally dependent on others, older adults may lose confidence in their ability to convince others with their lies. In an unpublished pioneer study (Elaad, 2016), surveying 39 residents (aged 66–94) of two Israeli nursing homes, low self-assessments of lie- and truth-telling abilities were reported for lie-telling and truth-telling abilities. Weak and dependent people who are in constant need of assistance may feel incapable of persuading others to believe them. However, caution is advised when using age to explain these results. To date, there is no evidence indicating that more independent older individuals will self-assess their lie- and truth-telling skills similarly.

In-Service Lie-Telling Experience

Agreeing to carry out a role or a job that demands the ability to tell lies better than the average person (e.g., interrogators, spies, attorneys, salespersons, and actors) may be related to relatively high a priori self-assessment of one's lie-telling ability. On the other hand, on-the-job-experience dealing with potential deception may enhance one's perception of being able to lie successfully. Table 23.1 shows that criminal interrogators' assessments of their lie-telling abilities are somewhat higher than the self-assessments of other groups, although differences did not reach statistical significance. Prosecutors, who were also expected to have relatively higher self-assessments of their lie-telling ability, in fact did not report such overestimated lie-telling ability.

Other Possible Mediators

No Reference to Other People

Lie- and truth-telling ability assessments are based on global impressions rather than on well-defined examples. Therefore, these assessments are prone to biases such as the "above average effect", which is a cognitive bias that causes people to overestimate their positive qualities and abilities and to underestimate their negative qualities. Therefore, one important question is, would the above-average effect hold when participants rate their lie- and truth-telling abilities with no reference to other people? Elaad (2015a) examined this exact question, using a sample of 60 students. Results indicated that the self-assessments of the lie-telling ability (Mean = 57.8, SD = 28.1) and truth-telling ability (Mean = 81.7, SD = 13.7) are somewhat higher than the average ability assessments obtained when reference to other people was made. However, the direction and the relative strengths of the assessed

abilities stay intact. Specifically, the truth-telling ability mean assessment is significantly higher than the lie-telling ability, $t_{(59)} = 6.63$, $p < .001$.

Perceived Importance

The different self-assessments of the lie- and truth-telling abilities may be linked to the importance people attribute to these abilities. Elaad (2015a) examined the importance mediator. Strong relations were found between the self-assessed abilities and their perceived importance. Specifically, the truth-telling ability received high importance ratings (Mean = 82.5, SD = 17.8), whereas the lie-telling ability received much lower importance ratings (Mean = 48.2, SD = 29.3). The difference was significant, $t_{(59)} = 7.73$, $p < .001$. People feel that it is more important to be a convincing truth teller than to be a good liar. The importance ratings can be considered in two ways: "it is important to possess a specific trait therefore I believe I possess it", or "I believe I possess the ability, therefore it is important to have it".

Test-Retest Reliability

Elaad and Sommerfeld (2016) provided insight into the reliability of the lie- and truth-telling abilities. They asked 100 students to self-assess their abilities twice in two occasions separated by 1–14 days. Test-retest correlations computed for the lie-telling ability, $r_{(100)} = .67$, and for the truth-telling ability, $r_{(100)} = .69$, were significant at the .001 level. It should be noted that the second ability assessment might have been affected by the experimental procedure that manipulated the feedback of the interrogator to participants' lies and truths, indicating that they were either believed or disbelieved. The reliability results may indicate that the lie- and truth-telling ability assessments are lasting attributes.

SELF-ASSESSED LIE- AND TRUTH-TELLING ABILITIES AND THE BIG FIVE PERSONALITY DIMENSIONS

Yet another question is whether people share similar notions about what constitutes positive and negative abilities. Some people associate the ability to lie successfully with dishonesty, a negative quality. Such people are expected to rate their ability to convincingly lie below average. Others may apply a double standard and consider their own lies to be an unavoidable necessity and therefore less innocuous than lies of other people (Bond & DePaulo, 2006). Yet other people may think that lying is a positive trait that serves them well in social interactions (Kashy & DePaulo, 1996). The latter are expected to rate their ability to lie successfully above average. With respect to truth-telling ability, some people self-assess it higher than others, mainly because people tend to believe them. Others may frequently encounter distrust when telling

the truth and may doubt their own truth-telling ability. It is suggested that personality attributes affect global lie- and truth-telling ability assessments.

Elaad and Reizer (2015) examined the contribution of personality dimensions to self-assessed lie- and truth-telling abilities using the Big Five personality inventory, one of the most dominant models of personality trait structure (McCrae & Costa, 1997). According to this model, five orthogonal dimensions capture the full range of human personality traits: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. Elaad and Reizer surveyed 174 Israeli undergraduate students using hierarchical regression models for predicting lie- and truth-telling ability assessments. They found that all Big Five dimensions contributed significantly to the self-assessments of the lie-telling ability. Specifically, higher levels of extraversion and openness to experience, and lower levels of agreeableness, conscientiousness, and neuroticism contributed significantly to higher assessments of the lie-telling ability. As to the ability to tell the truth convincingly, agreeableness contributed negatively, and extraversion contributed positively to higher truth-telling ability assessments.

A more recent study (Elaad, 2018) utilizing 160 Israeli participants drawn from a local community provided an opportunity to reexamine the contribution of the Big Five personality dimensions to high and low self-assessments of the lie- and truth-telling abilities. Using hierarchical regression models, lower levels of agreeableness contributed significantly to higher lie-telling ability assessments. All other dimensions failed to show any contribution to the self-assessed ability to tell credible lies. No Big Five personal dimensions predicted truth-telling assessments. Further, a recent unpublished study (Yaacov, 2017) surveying 192 local community members (108 females) examined the association between lie- and truth-telling abilities and the Big Five personality traits. Lower levels of agreeableness and higher levels of openness to experience contributed significantly to higher lie-telling ability assessments. Higher ratings of extraversion, openness to experience, and conscientiousness, and lower neuroticism ratings predicted higher truth-telling ability assessments.

These regression analyses indicate a robust association between self-assessed lie- and truth-telling abilities and the Big Five traits. To reach concise and more convincing conclusions, a mini meta-analysis (Goh et al., 2016) based on these three studies was conducted. The advantage of a mini meta-analysis over the individual regression analyses is the estimation of the overall effect size that places more weight on the reliability and the replicability of the findings than on individual effects that may or may not meet the adequate level of significance (Braver, Thoemmes, & Rosenthal, 2014). For the mini meta-analysis, Pearson correlations between lie-telling ability assessment and the Big Five dimensions in each study were calculated. The correlations are shown in Table 23.2. Similar correlations computed for the truth-telling ability assessments appear in Table 23.3.

Table 23.2 Data on correlations between lie-telling ability assessments and Big Five dimensions

	Agreeable	*Conscientious*	*Extraversion*	*Openness*	*Neuroticism*
Elaad and Reizer (2015); $N=174$	−.307***	−.112	.199**	.218**	−.107
Yaacov (2017); $N=191$	−.217**	−.105	.111	.143*	−.101
Elaad (2018); $N=160$	−.348**	−.095	.022	−.100	−.026
$M\,r_z$	−.296	−.104	.117	.156	−.080
$M\,r$	−.288	−.104	.117	.155	−.080
Combined Z	−6.52***	−1.65	2.06*	2.05*	−1.14

*$p < .05$, **$p < .01$, ***$p < .001$, two-tailed

$M\,r_z$=weighted mean correlation (Fisher's Z transformation for normalization). $M\,r$=weighted mean correlation (converted from r_z to r)

The combined Z value summarizes the Z (standard normal deviate) computed for each p-value with the appropriate sign. Calculations were performed according Stouffer's Z test, using the following formula (Goh et al., 2016)

$$Z_{\text{combined}} = \frac{\sum Z}{\sqrt{k}}$$

Where k refers to the number of independent Zs being combined

Table 23.3 Data on correlations between truth-telling ability assessments and Big Five dimensions

	Agreeable	*Conscientious*	*Extraversion*	*Openness*	*Neuroticism*
Elaad and Reizer (2015); $N=174$	−.229**	.190*	.473***	.205**	−.083
Yaacov (2017); $N=192$	.093	.171*	.368***	.242**	−.248**
Elaad (2018); $N=160$	−.104	.143	.110	.049	−.048
$M\,r_z$	−.075	.171	.345	.174	−.168
$M\,r$	−.075	.170	.332	.173	−.167
Combined Z	−1.70	3.37***	6.92***	3.21**	−2.19*

*$p < .05$, **$p < .01$, ***$p < .001$, two tailed

$M\,r_z$=weighted mean correlation (Fisher's Z transformation for normalization). $M\,r$=weighted mean correlation (converted from r_z to r)

Combined Z=Summarized Zs that correspond to each study's p-value

A separate mini meta-analysis was performed for each Big Five dimension across the three studies. Fixed effects were used, in which mean effect size (i.e., mean correlation) was weighted by sample size. All correlations were Fisher's z transformations for analyses, which were converted back to Pearson correlations for presentation (Goh et al., 2016). Across the three studies, the lie-telling assessments were positively associated with extraversion and openness to experience and negatively associated with agreeableness and neuroticism. Although consistent negative association emerged between conscientiousness and lie-telling ability assessments, the combined Z score was insignificant. Truth-telling assessments related positively with extraversion,

openness to experience, and conscientiousness and showed a negative relationship with neuroticism. Agreeableness showed inconsistent relations with truth-telling assessments, and the combined Z showed no significant effect.

These findings deserve further explanation. It seems that extroverts who are sociable, are more energetic, have positive emotions, tend to seek stimulation through social involvement, and are more talkative (McCrae & Costa, 1997), are drawn to social life, and have more opportunities to communicate with others. Compared with introverts who have fewer social interactions (Kashy & DePaulo, 1996), extroverts have more opportunities to tell lies and learn to be convincing when lying. Their frequent social interactions also contribute to their ability to be more convincing when telling the truth, and extroverts have been found to perceive themselves as good persuaders (Barrick & Mount, 1991).

Openness to experience also contributed to self-assessments of the ability to tell lies and truths convincingly. Openness to experience reflects one's degree of intellectual curiosity, independent thinking, creativity, and preference for novelty and variety (McCrae & Costa, 1997). Such individuals are curious about other people and spend time and effort collecting information from others. Confidence in their lie- and truth-telling abilities helps them in this matter. Furthermore, open-minded people who prefer a variety of experiences over a strict routine (Barrick & Mount, 1991) try to deceive more often than others, and learn how to become better liars. Specifically, experience with lying may be associated with higher self-assessments of lie-telling ability.

Agreeable people consider themselves as nice, friendly, and trustworthy (McCrae & Costa, 1997). Agreeableness is associated with the tendency to be genuine in one's relationships (Gillath, Sesko, Shaver, & Chun, 2010) and is linked with low self-assessments of the lie-telling ability, although it is not clear why agreeable people gave low ratings to their truth-telling ability (Elaad & Reizer, 2015).

Conscientiousness combines the tendency to show self-discipline, act dutifully, and focus on achievement. Conscientious people tend to plan rather than be spontaneous and are organized, dependable, responsible, and prudent. Conscientiousness shows negative associations with actual cheating (Day, Hudson, Dobies, & Waris, 2011) and positive association with honesty (Gillath et al., 2010). Conscientious people are driven by the desire to maintain an honest self-image and therefore rate high their truth-telling ability.

Neuroticism is described as having low self-confidence, pessimism, negative emotions, anxiety, and irritability. Neurotic people are unable to deal with stress and have difficulties controlling their impulses (McCrae & Costa, 1997). They tend to lie more than others (Conrads, Irlenbusch, & Rilke, 2013) and experience cognitive failure (Di Fabio & Palazzeschi, 2013), anxiety, and self-doubt. We may add that neuroticism is associated with low self-assessed lie- and truth-telling abilities.

It is evident that personality dimensions are associated with people's lie- and truth-telling ability assessments. Both low truth-telling ability raters and high lie-telling ability raters can now be described with reference to personality traits. Nevertheless, further investigation is warranted. For example, lifetime feedback on lies and truths that were believed or disbelieved may construct personality attributes. This may suggest that the self-assessed lie- and truth-telling abilities play a greater role in shaping personality than was previously thought.

LYING PREFERENCE AND SELF-ASSESSED LIE- AND TRUTH-TELLING ABILITIES

Yaacov (2017) examined the association between self-assessed lie- and truth-telling abilities and preferences for a deceptive option over a truthful one. Specifically, participants drawn from a local community were presented with four implausible scenarios of misconduct. They were then asked to simulate the role of the innocent respondent and defend themselves by providing a convincing story. Participants were presented with three optional stories: (a) telling a completely true but implausible story; (b) telling a story that was basically true although many implausible aspects of the event were removed from the description; and (c) telling a complete lie that makes sense.

Participants who provided high self-assessments of their lie-telling ability tended to choose deceptive alternatives more often than participants who provided low self-assessments of their lie-telling ability. Further, participants who were confident in their ability to tell the truth convincingly preferred the true but implausible option more than participants who rated low their truth-telling ability. This study contributed unique information about the association between lying preference and self-assessed lie- and truth-telling abilities.

SELF-ASSESSED LIE- AND TRUTH-TELLING ABILITIES AND PERFORMANCE IN THE CONCEALED INFORMATION TEST

The final section of this review offers evidence of association between participants' self-assessed lie- and truth-telling abilities and their responses to critical items in the Concealed Information Test (CIT). The CIT (Lykken, 1998) is an information detection test that consists of a series of multiple-choice questions, each containing mention of a single critical item that is related to the crime under investigation and several unrelated control items. It is assumed that a knowledgeable suspect would identify the critical item and respond to its appearance. Three physiological measures are commonly used to detect concealed knowledge—electrodermal responses, cardiovascular activity, and respiration changes.

Elaad and Sommerfeld (2016) recently examined the link between self-assessed lie- and truth-telling abilities and performance in one version of

the CIT (the Guilty Answer Test, GAT). The findings revealed that in the two-minute rest period before the test, skin conductance level (SCL) of guilty participants who had high assessments of their lie-telling ability was higher than that of lower lie-telling ability raters. Larger SCL responses may reflect arousal; therefore, it appears that guilty (but not informed, innocent participants) who had high assessments of their lie-telling ability were more aroused before the test compared with low raters. It seems that people who provide low self-assessments of their lie-telling ability realize that they are not fit to beat the polygraph and make no effort to change this outlook. Individuals who provide high self-assessments of their lie-telling ability may be motivated to prove their lie-telling skills and influence the polygraph outcome. This increased motivation is related to increased tonic SCL before the test and to enhanced SCRs to critical items during the test (e.g., Ben-Shakhar & Elaad, 2003). Lie-telling ability assessments were not associated with respiration or cardiovascular responses. Self-assessed truth-telling ability negatively contributed to the relative magnitude of skin conductance responses. In summary, individuals who provide high self-assessments of their lie-telling ability tend to elicit larger magnitude of electrodermal responses to critical CIT items.

DISCUSSION AND DIRECTIONS FOR FUTURE RESEARCH

Generally, people feel confident about their ability to tell the truth convincingly and expect others to believe them. This is particularly notable in situations involving personal social relations, including interactions within the context of law enforcement. Nevertheless, some people do not believe that they are convincing truth tellers. The low self-assessment of their truth-telling ability may be affected by their life experiences or moderated by factors such as apparent honesty. Some individuals have the advantage of an honest-looking face (i.e., baby face), while others are facially disadvantaged (Masip, Garrido, & Herrero, 2004). Life experience may teach people about their performance in truth-telling situations, which may partially explain truth-telling ability differences. However, this topic deserves further research.

Poorly self-assessed truth-telling ability may be damaging in criminal interrogations and in the courtroom. It may be speculated that people who have little confidence in their ability to tell the truth convincingly, may be more easily led to make false confessions. Police interrogators' tendency is to be suspicious (Elaad, 2003; Meissner & Kassin, 2002). This may even further erode the low confidence of these innocent suspects. Suspects may become frustrated and defensive, which in turn, further enhances interrogators' aggressive interrogation efforts (Kassin, 2005), paving the way to false confessions. Caution is advised when evaluating confessions obtained from suspects who have low evaluations of their own truth-telling ability.

The state of mind of distrusted truthful witnesses who lack confidence in their truth-telling ability is thought-provoking. Based on Bandura's (1997)

famous statement, "self-belief does not necessarily ensure success, but self-disbelief assuredly spawns failure" (p. 77), such a witness may fail to convincingly deliver the truth in the courtroom. Under cross-examination such witnesses might demonstrate low self-assurance, and their insecurity may convince the investigating attorney to adopt even more aggressive techniques. The ultimate result may be that a truthful but insecure witness will fail to convince the court. It is important to identify witnesses with low self-assessments of their truth-telling ability, as soon as possible, and provide appropriate support during the testimony.

As the present review suggests, personality attributes may play a role in low assessments of one's truth-telling ability. It was suggested that such people are socially insecure, reluctant to disclose feelings to others, and have difficulties in trusting other people. Neurotic features such as low self-confidence, pessimism, and negative emotions exacerbate such insecurity. Socially insecure individuals have difficulties dealing with stress and controlling their impulses. It may be hypothesized that low truth-telling ability raters tend to lie more than others and fail to deliver these lies successfully, which further contributes to their insecurity. These associations should be considered in future research.

The present review clearly shows that many people feel that convincing others to believe their lies is a difficult task. Social and religious norms may account for these feelings. Further research attention should be given to people who give extremely high assessments of their ability to lie convincingly. Future investigations might frame this topic around self-efficacy theory (Bandura, 1986), as higher self-efficacy levels in a specific area are related to better performance in that area.

Arguably, the higher self-assessment of one's ability to tell lies may be associated with better and more successful lying. Yaacov (2017) offered some support to this claim when reporting on the association between high lie-telling ability assessments and lying preferences. Schneider and Goffin (2012) found an association between the Perceived Ability to Deceive (PATD) scale scores and self-reported counterproductive workplace behaviors. Grieve and Hayes (2013) who investigated the relationship between PATD scores and actual ability to deceive successfully within the context of vocational faking simulations, found that PATD scores did not predict successful faking. Nevertheless, the relation between high confidence in one's lie-telling ability and actual lying behavior deserves further attention. Other related questions that warrant further investigation include: Do high lie-telling ability raters lie more frequently than others? In what situations would above-average lie-telling ability raters tend to use or refrain from using their supposed lying skills? These questions await answers using experimental methods.

High-raters of the ability to lie convincingly may be characterized by components of extraversion such as sociability and positive emotions. Further, high lie-telling ability raters scored high on elements of the openness to experience dimension. Attributes such as curiosity and independent thinking, emotional

intelligence, seeking novelty and variety, as well as spending time collecting information from other people, may be characteristic of this group of people.

It is possible that people who give high assessments to their lie-telling ability are inclined to tell lies more often than low lie-telling ability raters who are more introverted and have fewer opportunities for social interaction. The former, who are engaged in a variety of social experiences, gain more experience with lying, which promotes higher self-assessments of their lie-telling ability.

High lie-telling ability raters were found to be more responsive in the CIT than other participants. This finding raises the following question: Would these selected examinees engage more frequently in countermeasures, compared with other examinees when undergoing a polygraph test? The rationale underlying this notion is that people who believe that they are capable liars may also be more strongly motivated to prove the success of their deceptive skills in the polygraph test. People who have doubts about their truth-telling ability may feel that they are not going to be believed in any case. In the CIT context, they might use countermeasures to control items in which they are telling the truth. Clear answers to these questions may have important implications for the polygraph industry and procedures used in the context of law enforcement.

LIMITATIONS

The set of studies used in the present mini meta-analysis were conducted in Israel by a single research group. Specifically, all samples were comprised of Israeli (mostly Jewish) individuals. This is a primary limitation of the reported results, which impairs their external validity. In order to assess the role of self-assessed abilities in more general terms, it would be beneficial to conduct similar studies in different countries and societies, using a variety of individuals. The robust findings described in this review are replicable and should encourage further research that investigates the role of self-assessed lie- and truth-telling abilities in countries other than Israel. Such replications are expected to support the present outcomes.

Another limitation concerns the evaluation procedures used. The studies discussed in this chapter used a single question to define lie- and truth-telling ability assessments. Such a procedure may not preclude separate reliability tests for each study. It is advised to develop a questionnaire containing several items for each assessed ability to incorporate some aspects of that ability. An earlier attempt in this direction was the PATD scale (Schneider & Goffin, 2012), that was developed to examine individual faking differences in pre-employment testing. The PATD results offer some external support to the present outcomes. For example, PATD scores are negatively related to agreeableness and unrelated to conscientiousness.

Finally, social desirability and self-presentation may undermine the validity of participants' lie- and truth-telling ability assessments. Nevertheless, Williams and Gilovich (2008) showed that people truly believe in their

self-assessed ratings and take their estimates seriously enough to guide their actions. The multiple links between lie- and truth-telling ability ratings and demographic, personality, behavioral, and psychophysiological variables are sufficiently convincing to support the conclusion that most participants were sincere when they assessed their lie- and truth-telling abilities.

<div align="center">REFERENCES</div>

*References marked with an asterisk refer to studies included in the mini meta-analysis.

Ariely, D. (2012). *The (honest) truth about dishonesty.* New York: HarperCollins.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1992). Exercise of personal agency through the self-efficacy mechanisms. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of actions* (pp. 3–38). Washington, DC: Taylor & Francis.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1–26.

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology, 88,* 131–151.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234.

Braver, S. L., Thoemmes, F. J., & Rosenthal, R. (2014). Continuously cumulating meta-analysis and replicability. *Perspectives on Psychological Science, 9,* 333–342.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 3,* 203–242.

Conrads, J., Irlenbusch, B., & Rilke, R. M. (2013). Lying and team incentives. *Journal of Economic Psychology, 34,* 1–7.

Day, N. E., Hudson, D., Dobies, P. R., & Waris, R. (2011). Student or situation? Personality and classroom context as predictors of attitudes about business school cheating. *Social Psychology of Education, 14,* 261–282.

Debey, E., De Schryver, M., Logan, G. D., Suchotzki, K., & Verschuere, B. (2015). From junior to senior Pinocchio: A cross-sectional lifespan investigation of deception. *Acta Psychologica, 160,* 58–68.

Di Fabio, A., & Palazzeschi, L. (2013). Incremental variance in indecisiveness due to cognitive failure compared to fluid intelligence and personality traits. *Personality and Individual Differences, 54,* 261–265.

Elaad, E. (2003). Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology, 17,* 349–363.

*Elaad, E. (2006). How ideological crisis and prolonged external threat affect self-assessed abilities to tell and detect lies and truths. *Social Issues in Israel, 1*, 160–173 (Text in Hebrew).

*Elaad, E. (2009). Lie-detection biases among male police interrogators, prisoners and lay-persons. *Psychological Reports, 105*, 1047–1056.

*Elaad, E. (2015a). The distrusted truth: Examination of challenged perceptions and expectations. *Psychology, 6,* 560–571.

*Elaad, E. (2015b). Criminal prosecutors: Highly assessed lie-detection abilities and beliefs about defendants' deception. *Annals of Forensic Research and Analysis, 2*(2), 1017.

Elaad, E. (2016). Older adults' self-assessment of their abilities to tell and detect lies and truths: The case of occupants of protected nursing homes. Unpublished manuscript.

*Elaad, E. (2018). Personality, religiosity, and self-assessed abilities to tell and detect lies, tell truths, and believe others. *Advances in Applied Psychology, 3*(1), 1–9.

*Elaad, E., Lavy, S., Cohenca, D., Berholz, E., Thee, P., & Ben-Gigi, Y. (2012). Lies, truths, and attachment orientations in late adolescence. *Personality and Individual Differences, 52*(6), 670–673.

*Elaad, E., & Reizer, A. (2015). Personality correlates of the self-assessed abilities to tell and detect lies, tell truths and believe others. *Journal of Individual Differences, 36,* 163–169.

*Elaad, E., & Sommerfeld, E. (2016). Effects of guilt, disbelief, and assessed lie-truth telling abilities on physiological responses in the Guilty Action Test. *Psychology, 7,* 1075–1091.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116,* 429–456.

Gillath, O., Sesko, A. K., Shaver, P. R., & Chun, D. S. (2010). Attachment, authenticity, and honesty: Dispositional and experimentally induced security can reduce self-and other-deception. *Journal of Personality and Social Psychology, 98,* 841–855.

Gilovich, T., Savitsky, K., & Medvec, V. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology, 75,* 332–346.

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10,* 535–549.

Grieve, R., & Hayes, J. (2013). Does perceived ability to deceive = ability to deceive? Predictive validity of the perceived ability to deceive (PATD) scale. *Personality and Individual Differences, 54,* 311–314.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Kaplar, M. E., & Gordon, A. K. (2004). The enigma of altruistic lying: Perspective differences in what motivates and justifies lie telling within romantic relationships. *Personal Relationships, 11,* 489–507.

Kashy, D. A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology, 70,* 1037–1051.

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist, 60,* 215–228.

Lamarche, V. M., & Murray, S. L. (2014). Selectively myopic? Self-esteem and attentional bias in response to potential relationship threats. *Social Psychological & Personality Science, 5,* 786–795.

Lykken, D. T. (1998). *A Tremor in the blood: Uses and abuses of the lie detection* (2nd ed.). New York, NY: Plenum Trade.

Maner, J. K. (2014). Let's put our money where our mouth is: If authors are to change their ways, reviewers (and editors) must change with them. *Perspectives on Psychological Science, 9,* 343–351.

Masip, J., Garrido, E., & Herrero, C. (2004). Facial appearance and impressions of credibility: The effects of facial babyishness and age on person perception. *International Journal of Psychology, 39,* 276–289.

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist, 52,* 509–516.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26,* 469–480.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication.* Newbury Park, CA: Sage.

Morris, S. B., & DeShon, R. R. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups design. *Psychological Methods, 7,* 105–125.

Ruffman, T., Murray, J., Halberstadt, J., & Vater, T. (2012). Age-related differences in deception. *Psychology and Aging, 27,* 543–549.

Schneider, T. J., & Goffin, R. D. (2012). Perceived ability to deceive and incremental prediction in pre-employment personality testing. *Personality and Individual Differences, 52,* 806–811.

Spiro, R., Feltovich, P., Jacobson, M., & Coulson, R. (1992). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. Duffy & D. Jonassen (Eds.), *Constructivism and the technology of instruction* (pp. 57–75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester: Wiley.

Williams, L. A., & DeSteno, D. (2008). Pride and perseverance: The motivational role of pride. *Journal of Personality and Social Psychology, 94,* 1007–1017.

Williams, E. F., & Gilovich, T. (2008). Do people really believe they are above average? *Journal of Experimental Social Psychology, 44,* 1121–1128.

Yaacov, T. (2017). *Effects of plausibility, personality and self-assessed abilities to tell and detect lies and truths, on the tendency to prefer lies over truths* (Unpublished M.A. thesis). Department of Psychology, Ariel University.

"Passing" and the Politics of Deception: Transgender Bodies, Cisgender Aesthetics, and the Policing of Inconspicuous Marginal Identities

Thomas J. Billard

In May 2016, lawyers, academics, and activists gathered in London for the TransJustice conference, a workshop cosponsored by Birkbeck, University of London and City University London. The conference focused on legal issues facing transgender Britons, particularly in the domain of criminal justice. Among the issues discussed were the troubling implications of the 2003 Sexual Offences Act for transgender individuals. After extensive discussion and debate, the gathered experts reached the conclusion that, as currently written, the law could classify those who do not disclose their gender assigned at birth—or, as Goffman (1963) would put it, their *discreditable stigma*—prior to engaging in sexual intercourse as rapists (Fae, 2016; Sims, 2016). Thus, transgender individuals living out their authentic gender identities could be considered criminal deception when cisgender (i.e., non-transgender) individuals are not aware of the genders transgender people were assigned at birth.

This policy, while shocking in its own right, is merely reflective of broader cultural discourses about transgender identity and deception circulated in media narratives of transgender lives (Barker-Plummer, 2013; Halberstam, 2001; MacKenzie & Marcel, 2009; Sloop, 2000; Squires & Brouwer, 2002;

T. J. Billard (✉)
Annenberg School for Communication and Journalism,
University of Southern California, Los Angeles, CA, USA
e-mail: tbillard@usc.edu

Willox, 2003) and enacted in interpersonal interactions between cis- and transgender people every day (Schilt & Westbrook, 2009). These discourses position transgender people as deceivers who live out their genders for the purpose of seducing cisgender heterosexuals (Schilt & Westbrook, 2009; Squires & Brouwer, 2002), scrutinizing their appearances for signs of their "true" (i.e., assigned at birth) gender (Billard, 2016b; Rogers, 1992; Sloop, 2000). As such, these discourses delegitimate transgender identities by implicitly (and sometimes explicitly) suggesting that transgender identities are falsehoods that conceal the truth of "biological gender."

Central to these discourses of deception is the concept of "passing." Those transgender people who show no clear signs of the gender they were assigned at birth "pass" (as cisgender), while those who *do* show signs fail to "pass." Contradictorily, the successful attainment of cisgender aesthetics deemed "passing" legitimates a transgender person's claim to their gender identity (Billard, 2016b; Booth, 2015), but also renders them more malicious in their deception. In the words of Jack Halberstam (2001), the customary narrative of transgender life "recasts the act of passing as deception, dishonesty, and fraud" (p. 14). Consequently, discourses surrounding transgender people who "pass" justify punishment for their deception, whether through violence and murder (MacKenzie & Marcel, 2009) or through incarceration (as with the Sexual Offences Act).

Transgender passing thus raises important questions about the nature of *deception* and the status of *deceiver*, as well as about where authenticity and honesty diverge from one another. This chapter explores those tensions, challenging the application of the label of "deception" by the social majority to those of marginal identities, particularly inconspicuous ones, as it serves to delegitimate authentic identities and police the boundaries of social hierarchies. To do so, I first review the concept of "passing" as it has been articulated in both humanistic and social scientific literatures, before turning to scholarship on transgender passing in particular and the ways in which transgender identities are aesthetically evaluated. I then analyze media discourses of transgender deception and how these discourses legitimate anti-transgender violence. Finally, I discuss the ways in which the concept of deception serves to reinforce the marginality of subaltern identities more generally before concluding with implications of this argument.

CONCEPTUALIZING "PASSING"

Passing has been defined in widely varied ways, and from different disciplinary and ideological perspectives. While certain scholars argue that passing represents a rejection of socially imposed identities and the construction of new ones through constant performance (Caughie, 2005), others view passing as more utilitarian in function, namely to ensure the survival of the one who passes (Ahmed, 1999; Hobbs, 2014; Moriel, 2005). This latter

perspective further, and necessarily, implies that passing occurs only when someone from a marginal or oppressed social group crosses over into the dominant or privileged group (Moriel, 2005; Snorton, 2009). Other perspectives are less restricted, maintaining that passing challenges assumptions of immutable, physiologically based categories by demonstrating the insufficiencies of physiological evidence in accurate social categorization, regardless of the direction in which the passing occurs (Moynihan, 2010).[1] Yet each definition generally converges on a core notion of passing articulated quite clearly by Liora Moriel (2005): "a person from group A simply self-identifies as belonging to group B (and vice versa), is accepted as a member of the other group, and occupies that identity position without detection" (p. 177). It is in moments of detection, however, that discourses of deception arise.

The concept of passing in the US finds it origins in racial passing, and particularly in concerns about Black Americans passing undetected as White and/or freedmen in the eighteenth and nineteenth centuries (Hobbs, 2014). Cultural tensions around the status of race categorizations were high, as runaway slaves with light complexions evaded recapture and as rural Blacks moved into cosmopolitan environments to seek opportunities for economic and social advancement by passing for White (Hobbs, 2014). The resultant cultural panic manifested in legal battles, as courts sought to establish criteria by which to evaluate citizen's "official" racial identities (Gross, 1998). In the twentieth and twenty-first centuries, however, "passing" has been increasingly considered in contexts of class (e.g., Foster, 2005; Moriel, 2005), sexuality (e.g., Leary, 1999; McCune, 2014), and sex and gender (Caughie, 2005; Halberstam, 2001; Moynihan, 2010; Snorton, 2009; Squires & Brouwer, 2002).

Recent research on passing has occurred largely in the humanities, where it has been discussed in terms of performative identities and the transgression of social boundaries (Ahmed, 1999; see Ginsberg, 1996; McCune, 2014; Moriel, 2005; Rottenberg, 2003). As Sara Ahmed (1999) noted, this work generally positions passing as "a radical and transgressive practice that serves to destabilize and traverse the system of knowledge and vision upon which subjectivity and identity precariously rests" (p. 88). Moreover, as Elaine Ginsberg (1996) wrote, "passing is about identities: their creation or imposition, their adoption or rejection, their accompanying rewards or penalties" (p. 2). This line of inquiry has primarily focused on narratives, analyzing fictional texts as well as biographies for evidence of how social categories are constructed, enforced, challenged, and recuperated (Halberstam, 2001; McCune, 2014; Moriel, 2005; Moynihan, 2010; Rottenberg, 2003).

Within the social scientific literature, however, sociologists have developed a robust line of inquiry on passing, beginning with Erving Goffman's work on *stigma* and *stigma management*. From Goffman's (1963) perspective, passing is about "the management of undisclosed discrediting information about the self" (p. 42). That is, for an individual who has an invisible stigma known only to themselves (whether in the form of a disability or stigmatized

social identity, such as homosexuality), said individual "passes" when others lack any "discrediting information"—information that would reveal the stigma—about them. This passing may be incidental, rather than intentional, or it may be a deliberate strategy on the part of the individual who wishes to maintain their social status by concealing their stigma. In the case of the latter, Goffman discussed how stigmatized individuals work to control *social information*, whether that information is visual or behavioral, that might betray their stigma.

Thomas Kando (1972) applied Goffman's model of passing and stigma management to transsexuals, arguing that transsexualism is *discreditable*, rather than *discrediting*—which is to say, the stigma of transsexual identity can be discovered, but is not necessarily apparent without disclosure. As such, passing, as the state of non-discovery, can be seen as a method of stigma management (Kando, 1972, p. 477). Yet the transsexual's passing, as Kando noted, is not achieved through a singular act of (surgical) transition, but through a *continuing process* of gender. As Candace West and Don Zimmerman (1987) argued, gender is always a "routine, methodical, and recurring accomplishment" (p. 126), including in contexts of gender passing.

The truth of this argument is seen in one of the many points made by Suzanne Kessler and Wendy McKenna (1978) in their ethnomethodological investigation of gender: Genitalia, which we usually presume to be the site of an individual's gender, are generally concealed from public view in daily life, and yet we perceive people (often accurately) as being one of two genders based on external social cues, such as clothing, physical attributes, and comportment. In this sense, everyone "passes" as their gender, as the "truths" of our anatomies are concealed from others' inspection (Rogers, 1992; c.f. Zimmerman, 1992).[2]

In perhaps the most seminal investigation of transgender passing, Harold Garfinkel (1967) discussed how Agnes, a young transsexual woman, "achieved" her feminine gender—thus "passing"—through the management of social presentation and avoidance of situations which presented "the possibility of detection and ruin" (p. 137). While some scholars have focused their attention on Agnes' need to gain "cultural knowledge of how [womanhood] was to be done" (Wickes & Emmison, 2007, p. 314)—assuming she did not already have this knowledge—the instances of stigma management described by Garfinkel nearly all centered on her physical appearance and/or the physical manifestations of social performance. For example, Garfinkel (1967) recounted Agnes' concern over whether there would be an assured private space for her to change out of her wet bathing suit on a day at the beach with friends. In another instance, he described Agnes' initial false sexual modesty, which she performed to prevent her boyfriend from discovering she had a penis. Thus, the management of transgender stigma, even where social performances of gender are concerned, center in large part on visual components of passing.

Transgender Passing and Cisgender Aesthetics

Because transgender individuals' "passing" is fundamentally visual in nature,[3] and because passing more generally refers to an individual's undetected membership in a social group into which they were not assigned at birth, passing for transgender individuals necessarily implies the attainment of cisgender aesthetics. That is, for a transgender person to pass, they must appear to a stranger to "look cisgender." Such was the achievement of Agnes. That is not to say that a transgender person must necessarily appear conventionally attractive to pass, for, as Mary Rogers (1992) argued, "[t]hat Agnes apparently met cultural standards of female attractiveness is far less relevant than that she exhibited no physical characteristics visibly jeopardizing her appearance as a 'normal' female" (p. 182). Thus, the *acceptably gendered* appearance of a transgender individual is of central concern to their passing.

The attainment of cisgender aesthetics, deemed "passing," is not necessarily considered a desirable achievement, however. Rather, there is contentious debate among transgender theorists and activists about whether the desire to pass is "good" or "bad." In her argument for a shift toward "posttranssexualism," Sandy Stone (1994) wrote, "[t]he essence of transsexualism is the act of passing" (p. 168), but the rejection of passing—the deliberate choice to *not pass*—represents a more politically liberated transgender identity that we might call "posttranssexual." Likewise, Kate Bornstein (1995) argued that passing "becomes the outward manifestation of shame and capitulation. Passing becomes silence. Passing becomes invisibility. Passing becomes lies. Passing becomes self-denial" (p. 125). Yet, as Katrina Roen (2002) remarked, this perspective, while important for its points about regimes of gender enforcement, comes "perilously close to accusing passing transsexuals of having false consciousness" (p. 508). Indeed, passing is often considered a desirable achievement among transgender communities, while those who expend too little effort at passing are judged negatively (Roen, 2002). Moreover, for many transgender people, passing does not represent illusion or concealment, but self-actualization and psychic realness (Halberstam, 2001; Snorton, 2009).

But regardless of any moral judgments that could be debated concerning it, we must acknowledge the centrality of passing to transgender existence, particularly as mediated to the public. I have discussed elsewhere how news media representations of transgender individuals focus on the successful attainment of cisgender aesthetics in their discussions of transgender identity. As I argued, "Journalists weave aesthetic evaluations into their storytelling, using these evaluations to judge the legitimacy and the quality of the transgender person's identity" (Billard, 2016b). John Sloop (2000) argued similarly in his analysis of coverage of the murder of Brandon Teena, who journalists described as an "ideal man" because of his handsome "male" appearance. This reliance on aesthetic evaluations of transgender identity is

not unique to news media, however. For example, E. Tristan Booth (2015) found that in the narration of televised documentaries, the gender identities of transgender individuals were only acknowledged *after* surgical alteration— after their appearances had been made to conform to cisgender standards.

Even from a normative perspective, this reliance on cisgender aesthetic achievement as a marker of successful gender is problematic. As I have written previously,

> In evaluating transgender people aesthetically, journalists suggest to their read-
> ers that aesthetics are the primary determinant of the legitimacy and the qual-
> ity of a transgender person's gender identity, and reinforce the idea that gender
> identity is purely aesthetic, rather than a complex set of social characteristics and
> self-identifications. (Billard, 2016b)

Moreover, this reduction of the legitimacy and the quality of transgender identity to a transgender individual's success at passing necessarily sustains the assumption that cisgender identity is normatively "better," and that all other gender identities are mere facsimiles of the "natural" genders of cisgender men and women (Billard, 2016b).

Furthermore, in considering passing as a metric of successful gender achievement, we must consider the differences in the transition process between transgender men and transgender women. As Booth (2015) noted, "[b]ecause testosterone produces facial hair, trans men are usually perceived as male [i.e., pass] without facial surgery, whereas for trans women, estro-gen does not produce an equivalent marker of female-bodiedness, nor does it reduce one's height" (p. 124). Thus, the use of passing as criteria for authen-tic gender identity disproportionately affirms masculine gender identities while perpetuating the devaluation of feminine identities (see also Schilt & Westbrook, 2009).

MEDIA DISCOURSES OF DECEPTION
AND ANTI-TRANSGENDER VIOLENCE

Perhaps the most damaging cultural work performed by aesthetic evalu-ations of transgender identities is their supporting role in media discourses of deception. As many scholars have noted, passing is often seen as an act of deception—as the perpetration of an identity-based fraud (e.g., Caughie, 2005; Halberstam, 2001; McCune, 2014). Transgender individuals who pass are, in particular, often described as "double, duplicitous, deceptive" (Halberstam, 2001, p. 24), or as perpetrating a "'charade' or 'masquerade'" (Ginsberg, 1996, p. 16). Even where those labels have not been *explicitly* applied, as Sinéad Moynihan (2010) has remarked, "the metaphors of con-cealment, subterfuge and deception that have historically characterized pass-ing are still pervasive" (p. 2). This thus casts a transgender individual who

passes as "a predator who successfully preys on others by keeping them from the truth" (Sloop, 2000, p. 170).

Admittedly, transgender individuals who fail to pass cannot escape discourses of deception either. There is a passing double bind whereby the transgender person who passes is an insidiously successful deceiver, while the transgender person who does not pass is a monstrously unsuccessful deceiver. As Gordene MacKenzie and Mary Marcel (2009) noted, transgender women in particular who do not pass are depicted as "men in dresses" who deserve "disciplinary violence" and transgender women who do pass are depicted as tricksters who deserve the rage of their "victims" (p. 83).[4] Nonetheless, the "deception" of passing is by far the greater crime in these narratives, as transgender people who pass are regarded as far more insidious and discussed in far more defamatory ways.

A wide body of literature has documented discourses of deception in the coverage of transgender murder victims (e.g., Barker-Plummer, 2013; Bettcher, 2007; MacKenzie & Marcel, 2009; Schilt & Westbrook, 2009; Sloop, 2000; Squires & Brouwer, 2002; Willox, 2003).[5] Early work identifying these discourses focused on the murder of Brandon Teena, a transgender man who was murdered in 1993 by two male acquaintances who had discovered his sex assigned at birth, mere weeks after reporting that they had raped him (Sloop, 2000; Squires & Brouwer, 2002; Willox, 2003). Subsequent discourses in both mainstream and marginal media insisted Brandon was a woman and therefore a "lesbian deceiver" (Squires & Brouwer, 2002, p. 301) who "pass[ed] herself off as a boy" (Willox, 2003, p. 415). In the words of MacKenzie and Marcel (2009), applying this "deception narrative" to Brandon's story "privileges [his] female anatomy as the 'true' source of his gender identity, rather than his own consistent practice of living as a man and seeing himself as male" (p. 79), thereby delegitimating his claim to his masculine gender.

Bernadette Barker-Plummer (2013) noted similar narratives in coverage of the murder of Gwen Araujo, a feminine-passing non-binary transgender person (i.e., someone who identifies as neither male nor female) whom news media claimed had "tricked" hir[6] murderers into sexual contact by "pretending to be a woman whilst 'really' being a man" (p. 714; see also Bettcher, 2007).[7] Indeed, across a wide range of coverage of transgender murder victims, violence is identified as "a response to actual or perceived deception of the perpetrator by the transgender person" (Schilt & Westbrook, 2009, p. 446). As Barker-Plummer (2013) remarked, this causal attribution of violence to deception implies that "transgender identity is in itself a provocation" (p. 715). And this implication is drawn out more fully in these murderers' defense that they simply "panicked" at the realization they had been deceived (Barker-Plummer, 2013; Bettcher, 2007; MacKenzie & Marcel, 2009).

"Trans panic" defenses—the name given to the argument that a transgender person's murderer acted in the heat of passion at discovering their

presumed-cisgender sexual partner was actually transgender and should therefore not be held responsible for their crime—have been widely used in criminal cases (Lee & Kwan, 2014; Tilleman, 2010; Wodda & Panfil, 2015). However, these defenses have not only been deployed in the courtroom, but also echoed in media coverage of the cases (Barker-Plummer, 2013). In fact, Kristen Schilt and Laurel Westbrook (2009) found that 56% of stories covering the murder of transgender women employed the narrative of the panicked perpetrator. As they wrote, "[t]he belief that gender deception in a sexual relationship would result in fatal violence is so culturally resonant that, even in cases where there is evidence that the perpetrator knew the victim was transgender prior to the sexual act, many people involved in the case, including journalists and police officers, still use the deception frame" (Schilt & Westbrook, 2009, p. 457; see also MacKenzie & Marcel, 2009).

This defense of the violent perpetrators' actions in turn legitimates the use of violence against transgender people. Building off the narrative of deception, the trans panic defense suggests that because the transgender victim deliberately withheld their "true gender" (Schilt & Westbrook, 2009; Sloop, 2000), they were the true wrongdoer. In the words of Talia Mae Bettcher (2007), "victims of transphobic violence can be subject to blame shifting through accusations of deception" (p. 47). As such, the perpetrators are absolved (morally) of their crime and their use of violence is "justified" in such a way that anti-transgender violence at large is dismissed as *deserved*. As transgressors against truth through deception by successfully passing, transgender people, this discourse concludes, deserve their punishment.

AGAINST "DECEPTION"

The core function of these discourses of deception is the policing of identities—the maintenance of the boundaries that separate levels of social hierarchy and the delegitimation of claims to "new" identities that move individuals out of stigmatized ones. And, particularly in the transgender context, these discourses serve to insulate those cisgender individuals higher up in the social hierarchy from the "tainting" influence of transgender individuals' stigma. As Schilt and Westbrook (2009) discussed, framing transgender lives as "deceptive" works to protect the heterosexuality of the individuals who have sexual(ized) encounters with them; the man who slept with or hit on a transgender woman is truly straight because he was merely "deceived" by her passing into thinking she was a "natural" woman. Discourses of deception then excuse the man's "homosexual" behavior and blame (and justify punishment of) the transgender woman for dragging down his place in the social hierarchy.

While discourses of deception are particularly pertinent in the context of transgender passing, they do not operate solely in this sphere. Rather, analyzing how these discourses operate in the transgender context can further inform us about how discourses of deception operate in contexts of passing

more broadly. For instance, applying the label of "deception" to transgender passing serves to exclude transgender women from the broader category of womanhood and to exclude transgender men from the category of manhood. Whereas the act(s) of passing communicate clearly the transgender person's membership in their self-determined gender category, calling these acts "deceptive" recasts their membership as "trespassing." As Ginsberg (1996) wrote of racial passing, "[a]s the term metaphorically implies, such an individual crossed or passed through a racial line or boundary—indeed *trespassed*—to assume a new identity" (p. 3, emphasis in the original). Kimberlyn Leary's (1999) definition of passing as a marginalized person's masquerade performed "in order to enjoy the privileges afforded to the dominant group" (p. 85) further supports this notion. The passing person thus becomes an invader of the dominant group who passes into the group to exploit the benefits of membership.[8]

These narratives further work to cast the identities of passing people as inauthentic ones. People who pass are, in the words of Halberstam (2001), "excluded from the domain of the real" (p. 17), and their identities are rearticulated by those in the dominant group as mere appropriations of "realness." Aesthetic evaluations of passing play a role in this, as these evaluations contrast "attainment" against "aspiration," implicitly communicating that the "attained" identity is a replica (of whatever quality) of the "aspirational" identity in the same way that a replica Monet might be compared against its original (Billard, 2016b). When such comparisons circulate in media narratives of passing, passing identities will necessarily be represented as inauthentic performances of the social identities into which one has passed.

At their core, both in the transgender context and more broadly, discourses of deception, work to confine those who pass to their pre-passing state of discredited stigmatization. As Ginsberg (1996) neatly summarized it, passing permits stigmatized individuals to conceal their stigma, "escaping the insubordination and oppression accompanying one identity and accessing the privileges and status of the other" (p. 3). However, because, in the words of Squires and Brouwer (2002), "[d]ominant groups and institutions desire the ability to survey and evaluate all subordinates with ease, thereby ensuring knowledge and readiness" (p. 287), passing out of subordination must be punished. Such is the function of discourses of deception: public shaming and the justification of punishment through violence. As one salient example, Ahmed (1999) cited Nella Larsen's *Passing*, in which the main character, Clare, a black woman who passes as white, is "exposed" as a black woman and summarily killed. As Ahmed (1999) noted, "[s]uch a literalized punishment restores the narrative to its rightful order" (p. 91). Because a person with a stigmatized identity attempted to escape the plight of their stigma by passing, they were punished.

This then begs a final question of what a legitimating alternative discourse surrounding passing would be. I echo the sentiments of C. Riley Snorton (2009) in calling for interpretations of passing that focus not on the

challenges passing poses to dominant groups' mechanisms of social control or the maintenance of social hierarchies, but rather on the psychic role of passing and its power to enable passing individuals' self-actualization. In Snorton's (2009) words, "passing is not simply a question of how one is read but includes an agential power of affirming one's own reading of self. Definitions of passing therefore must also include its psychological function, that is, that it brings one's 'self' into view" (p. 87). Thus, in turning away from understandings of passing as a deceitful practice, we affirm the agency of those who pass to construct legitimate and authentic identities.

CONCLUSION

As I have demonstrated throughout this chapter, the label of "deception" is not a neutral marker of dishonesty or fraud. Rather, it is a socially fraught term used by the dominant group to discriminate between the legitimated and the delegitimated, the authenticated and the inauthenticated, the "righteous" and the "wicked." And as a power held by the dominant group, application of the label of "deception" serves to police the boundaries established around those lower in the social hierarchy, to maintain their subordination. In the context of transgender passing, discourses of "deception" serve to constrict transgender agency in self-identification, to delegitimate transgender identity claims, and to reinforce the stigmatization of transgender identities.

Media discourses of transgender life portray transgender individuals as deceivers whose identities are often ruses by which to "trick" cisgender heterosexuals into homosexual behavior (Barker-Plummer, 2013; Bettcher, 2007; MacKenzie & Marcel, 2009; Schilt & Westbrook, 2009). Yet, even in the few instances in which deception narratives do not focus on sexual enticement, these narratives still suggest that transgender individuals' identities are artificial and that their "dishonesty" reflects on their poor moral character (Sloop, 2000). These discourses of deception follow directly from instances of "passing," however, as individuals who "pass" as cisgender have "deceived" others into ignorance of their sex assigned at birth. While in other instances successful passing legitimates transgender individuals' claims to their gender identities (Billard, 2016b; Booth, 2015), where discourses of deception are circulated their passing becomes evidence of the insidiousness of their deception. And this insidiousness, in turn, justifies the (often murderous) violence committed against them (Bettcher, 2007).

As such, the label of "deception" and the discourses that surround its use delegitimate authentic transgender identities and regulate the divisions in social hierarchy between the transgender marginalized and the cisgender dominant. This regulation (re)establishes transgender individuals as "lesser" and ensures that they cannot escape their stigmatization by passing as a "natural" member of their self-identified gender. Moreover, such regulation criminalizes transgender passing in much the same way that the passing of

Blacks as White has been criminalized (Squires & Brouwer, 2002), and for much the same reasons: Members of the dominant group feel entitled to inconspicuous marginalized individuals' "discrediting stigma" so that they can enforce existing regimes of social control. Those who successfully conceal this stigma are thus punished for attempting to escape the marginalization of their stigma, which the dominant group regards as the natural and necessary state of the subjugated.

The ultimate consequence of these discourses of deception, as repeatedly mentioned, is the delegitimation of transgender identities. This delegitimation may manifest in social attitudes both toward transgender individuals (e.g., Nagoshi et al., 2008) and toward transgender rights (e.g., Miller et al., 2017), and it is apparent in public political debate over transgender policy (Billard, 2016a). Moreover, we see clear concretization of these discourses into social policy, such as in the Sexual Offences Act of 2003. We further see these discourses echoed in court cases (Lee & Kwan, 2014; Tilleman, 2010; Wodda & Panfil, 2015) and observe their impact on policing practices (Grant et al., 2011; James et al., 2016; Moran & Sharpe, 2004). Thus, the significance of these discourses of deception is not merely ideological, but practical, as they influence the state of transgender acceptance both socially and politically.

Finally, discourses of deception legitimate anti-transgender violence, which is rampant not only in Western contexts, but globally (Kidd & Witten, 2007; Stotzer, 2009). In her work on media violence, Sandra Ball-Rokeach (1971, 2001) has suggested that the greatest effect media has in the domain of violence is inducing the public to accept certain forms of violence as acceptable and legitimate (e.g., military, police) and reject others as unacceptable and illegitimate (e.g., civil unrest), which ultimately supports the status quo. As such, the great practical danger of how media discourses of transgender passing excuse anti-transgender violence is that it may produce in audiences the opinion that violence against transgender people should be accepted because it is justified. This consequently reifies a status quo in which the lives of transgender people not only have no value, but are also regarded as unworthy of the basic right to safety and security. While we currently lack empirical evidence to demonstrate this effect, discourses of deception surrounding transgender passing are firmly established and are likely operating already.

NOTES

1. It is important to note, however, as Ahmed (1999) did, that passing from a marginal group into the dominant group implies very different structural power relations than passing in the reverse direction.
2. The "passing" of cisgender individuals is of course, as Zimmerman (1992) noted in his critique of Rogers (1992), not *truly* passing, but merely "doing gender." However, the parallel between transgender passing and cisgender

"doing gender" productively highlights the ways in which transgender passing is constituted by the successful "doing" of gender.

3. Of course, other sensory elements of gender presentation and perception, such as vocal pitch and intonation (e.g., Hancock, Colton, & Douglas, 2014; Hancock & Garabedian, 2013; King, Brown, & McCrea, 2012), are key to transgender individuals' passing. However, discussion of these elements falls outside the scope of the present analysis, and their significance is less frequently discussed in both academic literature and popular media discourse than that of visual passing.

4. Although these discourses are, in many instances, applied to transgender men, it is worth noting that they are disproportionately applied to transgender women (MacKenzie & Marcel, 2009; Schilt & Westbrook, 2009; Westbrook & Schilt, 2014).

5. While news stories covering the murder of transgender individuals have been the subject of most media scholarship on transgender issues, such stories actually comprise only 14% of all national transgender news coverage (Billard, 2016c). However, these stories contain the most egregious forms of delegitimization of transgender issues and identities and are thus a subject of great academic interest. Moreover, these stories may, as MacKenzie and Marcel (2009) suggested, circulate more frequently in local, rather than national, transgender news coverage.

6. Hir is a gender-neutral pronoun and Gwen's preferred pronoun. See Barker-Plummer (2013) for a more thorough discussion of hir pronoun preferences.

7. Discourses of deception, while applied broadly to transgender people, are more frequently and more egregiously applied in *sexualized contexts*. As Schilt and Westbrook (2009) remarked, it is in sexualized spaces that gender identities become contentious, and cisgender individuals who are confronted with transgender identities in sexualized contexts deploy discourses that "regender" transgender people to the sex they were assigned at birth in order to protect and maintain the "heterosexual matrix" (pp. 450–451).

8. For an example of radical feminist work that advances this shocking notion in the context of transgender identity, see Raymond (1979). For a start to critiques of this argument, see Riddell (2006), Serano (2007), and Stone (1994).

References

Ahmed, S. (1999). "She'll wake up one of these days and find she's turned into a nigger": Passing through hybridity. *Theory, Culture & Society, 16*(2), 87–106.

Ball-Rokeach, S. J. (1971). The legitimation of violence. In J. F. Short Jr. & M. E. Wolfgang (Eds.), *Collective violence* (pp. 100–111). Chicago, IL: Aldine.

Ball-Rokeach, S. J. (2001). The politics of studying media violence: Reflections 30 years after the violence commission. *Mass Communication and Society, 4*(1), 3–18.

Barker-Plummer, B. (2013). Fixing Gwen: News and the mediation of (trans)gender challenges. *Feminist Media Studies, 13*(4), 710–724.

Bettcher, T. M. (2007). Evil deceivers and make-believers: On transphobic violence and the politics of illusion. *Hypatia, 22*(3), 43–65.

Billard, T. J. (2016a, September). *Constructing transgender identity in restroom-related policy debates.* Paper Presented at the Annual Meeting of the American Political Science Association, Philadelphia, PA.

Billard, T. J. (2016b, June). *The transgender body as artistic creation: Journalists' aesthetic evaluations of transgender identity.* Paper Presented at the Annual Meeting of the International Communication Association, Fukuoka, Japan.

Billard, T. J. (2016c). Writing in the margins: Mainstream news media representations of transgenderism. *International Journal of Communication, 10,* 4193–4218.

Booth, E. T. (2015). The provisional acknowledgment of identity claims in televised documentary. In L. G. Spencer & J. C. Capuzza (Eds.), *Transgender communication studies: Histories, trends, and trajectories* (pp. 111–126). Lanham, MD: Lexington Books.

Bornstein, K. (1995). *Gender outlaw: On men, women, and the rest of us.* New York, NY: Vintage.

Caughie, P. L. (2005). Passing as modernism. *Modernism/Modernity, 12*(3), 385–406.

Fae, J. (2016, June 1). Trans people face jail for rape if they have sex without sharing their gender history: Why even a kiss can see trans people branded as sex criminals. *Gay Star News.* Retrieved from http://gaystarnews.com/article/trans-people-face-jail-rape-sex-without-sharing-gender-history/.

Foster, G. A. (2005). *Class-passing: Social mobility in film and popular culture.* Carbondale, IL: Southern Illinois University Press.

Garfinkel, H. (1967). *Studies in ethnomethodology.* Eaglewood Cliffs, NJ: Prentice-Hall.

Ginsberg, E. K. (1996). The politics of passing. In E. K. Ginsberg (Ed.), *Passing and the fictions of identity* (pp. 1–18). Durham, NC: Duke University Press.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity.* New York, NY: Simon & Schuster.

Grant, J. M., Mottet, L. A., Tanis, J., Harrison, J., Herman, J. L., & Keisling, M. (2011). *Injustice at every turn: A report of the National Transgender Discrimination Survey.* Washington, DC: National Center for Transgender Equality and National Gay and Lesbian Task Force.

Gross, A. J. (1998). Litigating whiteness: Trials of racial determination in the nineteenth-century South. *Yale Law Journal, 108*(1), 109–188.

Halberstam, J. (2001). Telling tales: Brandon Teena, Billy Tipton, and transgender biography. In M. C. Sánchez & L. Schlossberg (Eds.), *Passing: Identity and interpretation in sexuality, race, and religion* (pp. 13–37). New York, NY: New York University Press.

Hancock, A., Colton, L., & Douglas, F. (2014). Intonation and gender perception: Applications for transgender speakers. *Journal of Voice, 28*(2), 203–209.

Hancock, A. B., & Garabedian, L. M. (2013). Transgender voice and communication treatment: A retrospective chart review of 25 cases. *International Journal of Language & Communication Disorders, 48*(1), 54–65.

Hobbs, A. (2014). *A chosen exile: A history of racial passing in American life.* Cambridge, MA: Harvard University Press.

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). *The report of the 2015 U.S. Transgender Survey.* Washington, DC: National Center for Transgender Equality.

Kando, T. (1972). Passing and stigma management: The case of the transsexual. *The Sociological Quarterly, 13*(4), 475–483.

Kessler, S., & McKenna, W. (1978). *Gender: An ethnomethodological approach.* New York, NY: Wiley.

Kidd, J. D., & Witten, T. M. (2007). Transgender and transsexual identities: The next strange fruit—hate crimes, violence and genocide against the global trans-communities. *Journal of Hate Studies, 6*(1), 31–64.

King, R. S., Brown, G. R., & McCrea, C. R. (2012). Voice parameters that result in identification or misidentification of biological gender in male-to-female transgender veterans. *International Journal of Transgenderism, 13*(3), 117–130.

Leary, K. (1999). Passing, posing, and "keeping it real." *Constellations, 6*(1), 85–96.

Lee, C., & Kwan, P. (2014). The trans panic defense: Masculinity, heteronormativity, and the murder of transgender women. *Hastings Law Journal, 66*(1), 77–132.

MacKenzie, G., & Marcel, M. (2009). Media coverage of the murder of U.S. transwomen of color. In L. M. Cuklanz & S. Moorti (Eds.), *Local violence, global media: Feminist analyses of gendered representations* (pp. 79–106). New York, NY: Peter Lang.

McCune, J. Q. (2014). *Sexual discretion: Black masculinity and the politics of passing.* Chicago, IL: University of Chicago Press.

Miller, P. R., Flores, A. R., Haider-Markel, D. P., Lewis, D. C., Tadlock, B. L., & Taylor, J. K. (2017). Transgender politics as body politics: Effects of disgust sensitivity and authoritarianism on transgender rights attitudes. *Politics, Groups, and Identities, 5*(1), 4–24.

Moran, L. J., & Sharpe, A. N. (2004). Violence, identity and policing: The case of violence against transgender people. *Criminology & Criminal Justice, 4*(4), 395–417.

Moriel, L. (2005). Passing and the performance of gender, race, and class acts: A theoretical framework. *Women & Performance, 15*(1), 167–210.

Moynihan, S. (2010). *Passing into the present: Contemporary American fiction of racial and gender passing.* Manchester, UK: Manchester University Press.

Nagoshi, J. L., Adams, K. A., Terrell, H. K., Hill, E. D., Brzuzy, S., & Nagoshi, C. T. (2008). Gender differences in correlates of homophobia and transphobia. *Sex Roles, 59*(7–8), 521–531.

Raymond, J. (1979). *The transsexual empire: The making of the she-male.* Boston, MA: Beacon Press.

Riddell, C. (2006). Divided sisterhood: A critical review of Janice Raymond's *The Transsexual Empire.* In S. Stryker & S. Whittle (Eds.), *The transgender studies reader* (pp. 144–158). New York, NY: Routledge.

Roen, K. (2002). "Either/or" and "both/neither": Discursive tensions in transgender politics. *Signs, 27*(2), 501–522.

Rogers, M. F. (1992). They were all passing: Agnes, Garfinkel, and company. *Gender & Society, 6*(2), 169–191.

Rottenberg, Catherine. (2003). Passing: Race, identification, and desire. *Criticism, 45*(4), 435–452.

Schilt, K., & Westbrook, L. (2009). Doing gender, doing heteronormativity: "Gender normals", transgender people, and the social maintenance of heterosexuality. *Gender & Society, 23*(4), 440–464.

Serano, J. (2007). *Whipping girl: A transsexual woman on sexism and the scapegoating of femininity.* Emeryville, CA: Seal Press.

Sims, A. (2016, June 11). Trans people could 'face rape charges' if they don't declare sexual history, warns trans activist. *Independent.* Retrieved from http://www.independent.co.uk/news/uk/home-news/trans-people-could-face-rape-charges-if-they-don-t-declare-sexual-history-warns-trans-activist-a7076546.html.

Sloop, J. M. (2000). Disciplining the transgendered: Brandon Teena, public representation, and normativity. *Western Journal of Communication, 64*(2), 165–189.

Snorton, C. R. (2009). "A new hope": The psychic life of passing. *Hypatia, 24*(3), 77–92.

Squires, C. R., & Brouwer, D. C. (2002). In/discernible bodies: The politics of passing in dominant and marginal media. *Critical Studies in Media Communication, 19*(3), 283–310.

Stone, S. (1994). The empire strikes back: A posttranssexual manifesto. *Camera Obscura, 29,* 150–176.

Stotzer, R. L. (2009). Violence against transgender people: A review of United States data. *Aggression and Violent Behavior, 14*(3), 170–179.

Tilleman, M. (2010). (Trans)forming the provocation defense. *Journal of Criminal Law and Criminology, 100*(4), 1659–1688.

West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & Society, 1*(2), 125–151.

Westbrook, L., & Schilt, K. (2014). Doing gender, determining gender: Transgender people, gender panics, and the maintenance of the sex/gender/sexuality system. *Gender & Society, 28*(1), 32–57.

Wickes, R., & Emmison, M. (2007). They are all "doing gender" but are they are all passing? A case study of the appropriation of a sociological concept. *The Sociological Review, 55*(2), 311–330.

Willox, A. (2003). Branding Teena: (Mis)representations in the media. *Sexualities, 6*(3–4), 407–425.

Wodda, A., & Panfil, V. R. (2015). Don't talk to me about deception: The necessary erosion of the trans* panic defense. *Albany Law Review, 78*(3), 927–971.

Zimmerman, D. H. (1992). They were all doing gender, but they weren't all passing: Comment on Rogers. *Gender & Society, 6*(2), 192–198.

"She Is My Roommate": Why and How Lesbian, Gay, and Bisexual Individuals Deceive Friends About Their Sexual Orientation

Yachao Li and Jennifer A. Samp

Same-sex romantic relationships are increasingly more socially acceptable and visible in the United States. According to a recent Gallup Poll (2017), the percentage of people who believe that same-sex couples should be recognized equally by the law has increased from 27% in 1996 to 61% in 2016. Even so, a same-sex relationship is still stigmatized (Gwartney & Schwartz, 2016; Herek, Chopp, & Strohl, 2007). Lesbian, gay, and bisexual (LGB) individuals[1] continue to be in the sexual minority and experience disapproval, prejudice, and discrimination based on their sexual orientation (Peplau & Fingerhut, 2007; Peplau & Spalding, 2000). Consequently, sexual minority individuals may intentionally pretend to be heterosexual during social interactions with others, a phenomenon we refer as "sexual orientation deception".[2] Deceiving others about one's minority sexual orientation is prevalent among the LGB community (e.g., Berg & Lien, 2009; Berger, 1990). For instance, Harrison (2003) reported that many adolescents feel

Y. Li (✉) · J. A. Samp
University of Georgia, Athens, GA, USA
e-mail: yachao.li@uga.edu

J. A. Samp
e-mail: jasamp@uga.edu

479

that they must pretend to be heterosexual to avoid rejection by their families and friends. Older LGB adults also purposely conceal their sexual orientation due to fear of mistreatment and discrimination (e.g., Choi & Meyer, 2016).

Although sexual orientation deception may protect LGB individuals from mental and physical abuse (e.g., Wells & Kline, 1987), it has several costs. Past research showed that gay men who pretend to be heterosexual show poorer social and intellectual adjustment (Braaten & Darling, 1965; Myrick, 1974) and experience greater depression, anxiety, and guilt (Berger, 1996). Recent studies on coming out have also found that sexual minority individuals who conceal their sexual orientation experience lower self-esteem, have greater psychological stress, and receive less social support; thus, they are more likely to suffer from mental health disorders, such as depression, suicide attempts, and substance use and addictions (Morris, Waldo, & Rothblum, 2001; Rosario, Hunter, Maguen, Gwadz, & Smith, 2001). In addition, the concealment of sexual orientation may also have adverse effects on physical health. For instance, closeted sexual minority individuals are more likely to engage in unprotected sex and substance abuse (Rosario et al., 2001) and are therefore susceptible to develop cancers and other infectious diseases (Cole, Kemeny, Taylor, & Visscher, 1996). Thus, an investigation of sexual orientation deception is warranted.

Deception attempts may have different targets. This chapter focuses on deceiving *friends* about one's sexual orientation. Compared to heterosexual individuals, sexual minority people report receiving a greater amount of social support from their friends (Detrie & Lease, 2007; Ueno, 2005). Social support from friends helps LGB individuals foster a sense of identity (Kertzner, Meyer, Frost, & Stirratt, 2009; Nardi, 1999; Stanley, 1996), reduces their stress of being a sexual minority (Frost & Meyer, 2009; Meyer, 2003), and mitigates the risk for developing adverse mental health outcomes (Bagwell et al., 2005; Lyubomirsky, Tkach, & DiMatteo, 2006). Thus, friendships are important interpersonal relationships for LGB people. However, many LGB individuals report that in order to maintain their friendships, they have pretended to be heterosexual in front of their friends due to the fear of rejection, betrayal, and even violence (e.g., Diaz, Ayala, Bein, Henne, & Marin, 2001; McLean, 2001).

Existing research has paid less attention to deceiving friends about sexual orientation; we know little about why and how LGB people pretend to be heterosexual in front of their friends. Thus, this chapter examines the motivations and strategies of people's sexual orientation deception toward friends. We first report a study that identifies factors influencing individuals' decisions to deceive or tell the truth to their friends about their sexual orientation. Then, we summarize people's communicative strategies of those deception attempts. Finally, we link these findings to theories that explain and predict how individual and cognitive factors affect people's communicative behaviors.

METHOD

Participants

Eighty-seven individuals were recruited through Amazon's Mechanical Turk.[3] The eligible criteria were (a) being 18 years or older, (b) self-identifying as lesbian, gay, or bisexual, and (c) having friends in real life. Twenty-one participants (24.1%) self-identified as lesbian, 27 (31.0%) people reported they were gay, and 39 respondents (19 males, 20 females) identified themselves as bisexual. In total, there were slightly more males (54.0%) than females. Participants ranged in age from 18 to 52 years old, with a mean age of 31.17 years (SD = 26.87). The majority identified themselves as Caucasian (44.8%) and Asian (41.4%), with the remainder being Hispanic (6.9%), African American (3.4%), and other (3.4%).

Procedures

After providing informed consent, participants were first instructed to think about three friends in real life and list their initials or nicknames on the online questionnaire. This task helped promote reflection on deception attempts with specific friends. All participants were able to identify three friends. Then, respondents were asked whether they had pretended to be heterosexual in front of any of the friends they just specified (58.5% answered "yes"). Next, participants responded to open-ended questions regarding the reasons for their choices. For participants who had pretended to be heterosexual, they also reported their communicative behaviors used in their deception attempts. Finally, participants reported their own demographic information, as well as their friends' age, gender, and sexual orientation.

Data Analysis

Content analyses were conducted to inductively identify themes present in the data (Neuendorf, 2002). The data from open-ended questions were first coded into discrete thought units. Each discrete thought unit represented an independent and complete idea, which can be reflected in a part of a sentence, in a single sentence, or in multiple sentences. From the 51 participants who pretended to be heterosexual, we identified 209 discrete thought units of the motives for sexual orientation deception and 87 discrete thought units of the communicative behaviors during these deception attempts. From the 36 participants who reported having not deceived their friends, 54 discrete thought units were identified for why they did not deceive. All discrete thought units were coded by a primary research assistant. A second coder coded a randomly drawn sample of 20% responses for reliability purposes. Inter-coder reliability was high (Cohen's Kappa = 1.0).

Once discrete thought units were identified, each unit was coded using a standard thematic analysis approach (Braun & Clarke, 2006). This data-driven approach allowed response categories and themes to emerge from the participants' responses rather than prior conceptual categories (Boyatzis, 1998). A constant comparative process (Corbin & Strauss, 1990) was utilized to compare the emerging themes until saturation was reached. At that point, no new categories emerged and all discrete thought units fit within a theme. The findings were offered to several research participants and LGB individuals not in this study to gain a sense of whether the themes represented their personal experiences. The member checks affirmed the analysis.

Friend Profile

All participants listed three friends and provided the friends' age, gender (1 = female, 2 = male), and sexual orientation (1 = lesbian, gay, or bisexual, 2 = heterosexual). Because participants reported their sexual orientation deception based upon those friends, we averaged the three friends' age, gender, and sexual orientation, respectively, to create an overall friend profile.[4] For instance, Rebecca listed three friends: Ryan (28, male, gay), Kyle (26, male, bisexual), and Nancy (36, female, heterosexual). Thus, we considered Rebecca having "a friend" who was 30 years old, "slightly male" (gender mean = 1.67), and "slightly non-heterosexual" (sexual orientation mean = 1.33). By creating these friend profiles, we were able to utilize independent samples *t*-tests to examine whether friend characteristics differ between those who engage in deception attempts and those who did not. On average, the "averaged friend" was 28.51 years old (SD = 8.47), "slightly male" (*M* = 1.46, SD = .42), and "slightly heterosexual" (*M* = 1.67, SD = .42).

RESULTS

Quantitative Analysis

First, an independent samples *t*-test explored whether the occurrence of sexual orientation deception (0 = no deception, 1 = deception) differed between genders. Results showed that male participants reported greater sexual orientation deception (*M* = .72, SD = .45) than female respondents (*M* = .43, SD = .50), $t(79) = -2.90$, $p = .005$. In addition, a Pearson chi-square test affirmed that there were more males than females who pretended to be heterosexual in front of their friends, $\lambda^2 (1) = 7.93$, $p = .005$ (see Table 25.1).

Second, a one-way ANOVA examined whether groups with different sexual orientations engaged in different deception attempts toward their friends. Results indicated significant group differences, $F(2, 84) = 15.20$,

Table 25.1 Crosstab between sexual orientation deception and participants' gender and sexual orientation

Have you ever pretended to be heterosexual?	Gender		Sexual orientation		
	Male (n = 47)	Female (n = 40)	Lesbian (n = 21)	Gay (n = 27)	Bisexual (n = 39)
Yes (n = 51)	34	17	3	18	30
No (n = 36)	13	23	18	9	9

Note Pearson chi-square tests (see text) of crosstabulation showed that there were more males than females who pretended to be heterosexual, and fewer lesbian participants engaged in sexual orientation deception than gay and bisexual respondents

$p < .001$. Post hoc analyses using Bonferroni correction showed that bisexual ($M = .77$, SD $= .43$) and gay ($M = .67$, SD $= .48$) participants were significantly more likely than lesbian respondents ($M = .14$, SD $= .36$) to report sexual orientation deception (ps $< .001$). Also, a Pearson chi-square test affirmed that fewer lesbian participants pretended to be heterosexual than gay and bisexual respondents, χ^2 (2) $= 23.12$, $p < .001$ (see Table 25.1).

Third, an independent samples t-test compared age differences between those who pretended to be heterosexual and those who did not. The result indicated that participants who engaged in sexual orientation deception were older ($M = 32.94$, SD $= 9.38$) than those who did not deceive their friends ($M = 28.67$, SD $= 7.24$), $t(85) = -2.40$, $p = .024$.

Last, several independent samples t-tests compared "friend profile" (friends' age, gender [$1 =$ female, $2 =$ male], and sexual orientation [$1 =$ lesbian, gay, and bisexual; $2 =$ heterosexual]) differences between those who engaged in sexual orientation deception and those who did not. Results showed that friends being deceived by participants were older ($M = 31.67$, SD $= 9.41$) than those who were not deceived ($M = 28.67$, SD $= 7.24$), $t(85) = 3.06$, $p = .003$. They also tended to be "more heterosexual" ($M = 1.76$, SD $= .36$) than those who knew participants' sexual orientation ($M = 1.52$, SD $= .47$), $t(63) = 2.55$, $p = .013$. No friend gender differences were observed between those who pretended and those who did not, $t(60) = 1.22$, $p = .228$. Thus, the friends in front of whom participants pretended to be heterosexual were older and more likely to be heterosexual, regardless of their gender.

Qualitative Analysis

Participants reported why they decided to deceive or tell the truth to their friends about their sexual orientation. Among respondents who pretended to be heterosexual ($n = 51$, 58.6%), they also indicated what they said and did during those deception attempts. Table 25.2 includes the percentage of discrete thought units that were coded into each theme and category.

Table 25.2   Themes emerged from participant responses

Themes	Percent of discrete thought units (%)
Reasons to deceive ($n = 209$)	
a. Friend characteristics	10
b. Fear of negative consequences	63
b.1. Relational consequences	22
b.2. Personal consequences	41
c. Homophobia	24
c.1. Friends' homophobia	18
c.2. Internalized homophobia	6
d. Lack of communication efficacy	7
Reasons not to deceive ($n = 54$)	
a. Friend characteristics	6
b. Identity acceptance	70
c. Friend support	24
Communicative strategies to deceive ($n = 87$)	
a. Individual strategy	79
a.1. Passive strategy	26
a.2. Active strategy	46
a.3. Interactive strategy	7
b. Couple strategy	21
b.1. Avoidance or denial	10
b.2. Naming	11

Note The number of discrete thought units representing reasons to deceive, reasons not to deceive, and communicative strategies to deceive were 209, 54, and 87, respectively. Percentages indicate number of discrete thought units that fit within each theme and category

The results presented herein rely upon direct quotes from the participants' responses in order to give priority to their voices. The examples chosen are representative of the responses of participants and include the necessary background information of the deceptive attempts. Any names are pseudonyms to protect confidentiality. Notes in parentheses are added by the authors. Notes in square brackets are kept from participants' original responses.

Reasons to deceive. The first research goal was to examine *why* sexual minority individuals engage in sexual orientation deception. Four themes emerged among motives for those deception attempts. The first theme relates to evaluations of the friends' characteristics, such as their age, religious beliefs, and sexual orientation. The second theme relates to appraisals of negative consequences of revealing one's sexual orientation. The third theme focuses on friends' homophobia and participants' internalized homophobia that encourage the deception attempts. The last theme represents whether people believe they have the abilities to communicate about minority sexual orientation with their friends.

Friend characteristics. The first theme in participants' reasons to engage in sexual orientation deception had to do with their evaluations of their friends. Some believed that because of their friends' certain demographic

characteristics, pretending to be heterosexual was appropriate and apparent. For instance, Sarah (35, lesbian) stated, "In the case of Tom, he is an older gentleman. I don't want to confuse him." David (47, gay) made a similar comment, "This (coming out) only occurred when we were all younger, around 13 years old." These examples echo the quantitative results indicating that when a friend was considered older, people tended to deceive him or her about their minority sexual orientation. In addition, some participants referred to their friends as "very religious" and thus wanted to behave in a way that matched their expectation: "My friends are religious in such a way that their religion prevents them from accepting GLBTQ" (Justin, 25, bisexual). Other participants mentioned that pretending to be heterosexual is taken-for-granted when their friends are heterosexual: "Because they are all heterosexual, for sure I act straight in front of them" (Jasper, 24, gay). A similar statement was from Mollie (20, lesbian), "Because they are hetero(sexual), that's why." These comments are consistent with the quantitative results that when a friend was heterosexual, LGB individuals were more likely to deceive about their sexual orientation.

Fear of negative consequences. The second theme that emerged from participants' motives for sexual orientation deception focused on appraisals of negative consequences resulting from coming out. First, participants indicated that if they were honest about their sexual orientation, they would hurt or lose the friendships. Debra (36, lesbian) said, "Ken has been my friend for years and has made inaccurate comments about the LGBT community, so I did not want to lose his friendship." Jimmy (30, bisexual) expressed similar concerns, "If I told them I had feelings about men, they might think I was hitting on them and our relationship was over."

In addition, participants also worried about negative personal consequences of coming out and thus decided to pretend to be heterosexual. Some feared negative opinions and judgment from others, for instance, "I feel like they will change their minds about me. Have a negative opinion" (Kim, 27, bisexual), "I fear my friends might have biases, consciously or unconsciously, against me and minority sexualities" (Steven, 22, gay), and "I felt like they would either make fun of me, judge me, or not want to talk to me at all anymore" (Jackson, 28, gay). Other participants showed concerns about privacy management: "My friends are close and in touch with members of my own family and I am afraid it (the participant's sexual orientation) will become common knowledge because I am not out to my families" (David, 47, gay). Bucky (31, bisexual) stated, "Because I hadn't come out yet to anyone and I am scared of the possibility that people might find out." Thus, for David, Bucky, and others alike, negative personal consequences also mean losing control over their private information and a greater risk for sexual orientation revelation.

Homophobia. The third theme that emerged among participants' motives for sexual orientation deception represented negative attitudes and

feelings toward homosexuality and sexual minority people. First, participants described their friends' homophobia that they learned from previous interactions as reasons for sexual orientation deception. Laura (27, bisexual) stated, "I believe they won't be accepting and even hate LGBT because they make remarks all the time." Andrew (30, gay) provided more details about how he learned of his friends' anti-LGB opinions:

> I know how they feel about people that are gay, lesbian, or bisexual… and it is not positive. I remember watching a music video with them and it showed a gay couple [males] kissing on a beach, and smiling at each other because they were so happy and content. And I remember how happy it made me to watch it, but how they had said it was gross. I will never forget it because it really disappointed me and made me feel disconnected from both of them.

Some participants also referred to their friends' homophobia as reasons that would lead to negative consequences of coming out. Amanda (27, bisexual) said, "I fear that they would abandon me if I were honest because they think of us as some shit people." In addition, several participants demonstrated high levels of internalized homophobia or LGB individuals' personal acceptance and endorsement of societal prejudice and stigma toward sexual minorities (Herek, Gillis, & Cogan, 2009). Angie (26, bisexual) wrote, "Because sharing with them as I am bisexual make(s) me feel ashamed." Zac (42, gay) also said, "I want to be honest but I am ashamed of being gay. I have conflicted feelings… I eventually decided to be straight." Thus, for Angie and Zac, lack of self-acceptance as sexual minorities motivated their sexual orientation deception.

Lack of communication efficacy. The final theme that emerged from participants' reasons to engage in sexual orientation deception focused on people's communicative abilities to talk about minority sexual orientation. Bucky (31, bisexual) stated, "I didn't really know how to broach that kind of topic with people so I have to pretend to be heterosexual in front of my friends." Amy (30, bisexual) said, "How can you bring up the conversation? Like 'hi, by the way, I am bi(sexual).' That's awkward." Here, Bucky and Amy showed that they feel unable or uncomfortable to talk about their sexual orientation, which encourages them to choose the "easier" option: "pretending to be straight is much easier. People don't need to say, 'I am heterosexual'" (Lee, 29, gay).

Reasons not to deceive. To better understand individuals' motives for sexual orientation deception, we asked those who were honest about their sexual orientation, "why did you decide not to pretend to be heterosexual?" Here, we coded the participants' responses into three themes. The first relates to friend characteristics, primarily their sexual orientation. The second theme focuses on participants' high self-acceptance of their sexual orientation. The last theme represents the perception that their friends are supportive.

Friend characteristics. Like those who pretended to be heterosexual, respondents decided to be honest about their sexual orientation based upon their evaluations of their friends. Most participants attributed their disclosure decisions to the fact that their friends were also not heterosexual: "My friends [including A, B, and C] are not heterosexual themselves, so there is no point to hide my sexuality from them" (Martin, 25, gay). Lucy (34, lesbian) provided a more detailed story.

> There has never been any reason to hide my sexual orientation from any of these three people. One, I met at work, and while she is closeted [and bisexual] she has good gaydar (a slang word meaning the ability to assess others' minority sexual orientation) so she figured me out anyway right away. Another, we were introduced by a mutual friend who knew both of us were lesbians. The third person, I met through feminist politics, and she is lesbian too.

Along with previous responses that people pretended to be heterosexual because their friends were heterosexual, Martin and Lucy's comments show that friends' sexual orientation (heterosexual vs. non-heterosexual) plays a role in people's decision to deceive about their minority sexual orientation. Indeed, the *t*-test also indicated that the friends being deceived were more likely to be heterosexual.

Identity acceptance. The second theme that emerged among participants' motives for sexual orientation honesty was the strong self-acceptance of one's identity. Indeed, most participants who did not pretend to be heterosexual expressed their comfort and pride of being LGB. For example, John (34, gay) wrote, "I don't want to fake my sexual identity. I am what I am. I am comfortable with it and share it with my friends." Cindy (20, lesbian) commented, "I like what I am and my authentic self. I do not want to hide it from anyone." Linking these responses to the "internalized homophobia" category of reasons to deceive, the degree to which an individual accepts his or her sexual orientation influences their decisions about sexual orientation deception.

Friend support. The last theme that emerged from respondents' discussions of their reasons not to deceive represented the opinion that "true friends don't judge" (Annalisa, 21, lesbian). Participants indicated that a friend should behave in ways that convey understanding, acceptance, and support for their sexual orientation. For example, one participant stated, "If they were friends, they would accept me as I am" (Lucy, 34, lesbian). Another participant indicated, "I don't just friend anyone. I hang out with them several times before I come out [so I know if they can be my friend]. If they accept we will be friends. If they don't I will just move on" (Mary, 28, lesbian). Here, Mary utilized friends' support as a criterion to develop and maintain a friendship. In other words, unlike those who fear of betrayal or rejection by friends, participants feel that they have control over their friendships. Elli (25, lesbian) elaborated this point, "I can choose whom will be my friend, the one who accepts me."

Communicative strategies to deceive. The second research objective of the study was to explore communicative strategies used for sexual orientation deception. Two themes emerged among the communicative behaviors during the deception attempts. The first theme represented strategies used by most sexual minority individuals to "cover" themselves, including three categories: passive, active, and interactive strategies. The second theme focused on how LGB individuals in same-sex relationships pretended to be heterosexual when their romantic partners are present in their daily life, including avoidance or denial, and naming.

Individual strategy. First, participants indicated that because heterosexuality is taken-for-granted in society, they could pretend to be heterosexual by "doing nothing" or "keep(ing) silent." Steven (22, gay) mentioned, "I did not say when I felt attraction to the same-sex as they would when they would say they were attracted to a person of the opposite sex." Brandon (25, gay) said, "I had a girlfriend before and I never uttered a word about what I thought about other boys. No one could ever know I am gay." Jimmy (30, bisexual) also utilized this strategy when his friends made inaccurate comments about LGB people, "I just did not speak up about the inaccuracy of his statements and instead let him continue to talk, neither agreeing or disagreeing." Here, sexual minority individuals "took advantage of (the fact) that most people are straight" (Jason, 37, gay), and pretended to be heterosexual by hiding their sexual orientation.

The second category of individual strategies represents the situation when participants actively talked about heterosexual topics with friends. Most gay participants mentioned that they "made comments about women" (Mark, 37, gay), "verbally expressed my interest in girls" (Jackson, 28, gay), and "involve in their (friends') locker room talk when they are chatting" (Lee, 29, gay). "To make it more real," Billy (19, gay) said, "I made improper, very bad jokes about women." Moreover, some gay participants talked about other "straight topics" like "sports and politics" (Jasper, 24, gay) with their friends. While all discrete thought units that fit into this category were from the study's gay participants, a lesbian participant conducting the member check commented, "this (talking about men) is not uncommon among lesbians."

Third, some gay and lesbian participants indicated that they engaged in direct interactions with someone of the opposite sex, a strategy we refer as an interactive strategy. Those interactions included heterosexual dating (e.g., "I dated guys." Sarah, 35, lesbian) and flirting. For instance, Wade (24, gay) wrote, "I tried hitting on girls around them so it seemed like I was interested in girls." Randy (26, gay) also commented that "I stock girls in front of them". In addition, some lesbian and gay participants also engaged in heterosexual intercourse. Lilly (20, lesbian) said, "I had sex with some fraternity guys and you know the girls (her friends) will know. No one doubted me." In one case, a gay participant married a woman and raised children:

"I am a married man and had three boys. They never [even my wife and children] asked (about my sexuality)." Notably, almost all participants who engaged in the interactive strategy for sexual orientation deception had pointed out some negative consequences of this tactic. For example, Lily said, "I felt guilty to myself for lying and taking advantage of the guys." The married gay participant stated, "I'm ashamed of myself... I have to lie and constantly monitor my behaviors so as to conform, which takes quite a toll on my psyche."

Couple strategy. The second theme emerged from responses of participants who were in same-sex relationships. First, participants showed that they either avoided introducing their partners in front of their friends or denied that they had a partner when asked by peers. Julia (29, lesbian) said, "I just never talked about the girlfriend that I was with at the time because I knew how they would feel about it." In addition to topic avoidance, some participants tried to avoid the situation where partners were invited. One participant indicated, "If I could, I found any kinds of excuses to avoid his parties [he asked me to bring my girlfriend]" (Kevin, 39, bisexual dating another man). Joe (25, gay) provided an example of denial, "I pretended I am single, but I am with another guy." Similar to but different from the passive strategy, participants hid their same-sex partners to conceal their own sexual orientation and eventually pretended to be heterosexual.

In addition, other participants acknowledged that there was a "special" person in their life but created a different name for their romantic partner. For instance, Julia (29, lesbian) called her girlfriend "a good friend," and Angie (26, bisexual) and Olaf (23, gay) referred to their partners as "my roommate." Peter (36, gay) explained why he named his boyfriend as a "cousin": "Though coming out to my parents was difficult [harder for them than me], and it took my mother years to fully accept George (Peter's boyfriend). We decided to keep this inside the house and told others George was my cousin." Tina (32, bisexual) remarked on the benefits of naming, "Grace (Tina's girlfriend) was happy because she got to know my friends and I was happy because I was able to remain closeted by calling her my 'best friend since high school.'" For Tina, naming helps her include the partner in her social network while retaining some levels of privacy. In one case, Henry (30, gay) also took advantage of his partner's bisexual name, "His name was Chris... I told my friends Chris was my partner and blah blah blah... they thought Chris was a girl." Here, a partner's gender-neutral name was utilized to help cover one's sexual orientation.

In summary, individual and friend characteristics (e.g., age, gender, sexual orientation, attitude toward homosexuality) and cognitive appraisals of coming out conversations (e.g., fear of negative consequences and communication efficacy) appear to determine whether LGB individuals deceive friends about their sexual orientation. When doing so, single and coupled sexual minority people utilized a variety of strategies, including taking advantage of

the heterosexual assumption, talking about and directly interacting with persons of the opposite sex, and avoiding or denying and renaming their same-sex partners.

DISCUSSION

This chapter offers insights into *why* and *how* some LGB individuals pretend to be heterosexual in front of their friends. Our data showed that individuals were likely to deceive their friends about their sexual orientation when they (a) considered their friends as older, religious, and heterosexual; (b) feared of negative relational and personal consequences of coming out; (c) perceived their friends as homophobic or had internalized stigmas toward themselves; and (d) believed a lack of communicative ability to talk about sexual orientation. In contrast, LGB people were likely to be honest about their sexual orientation when they (a) interacted with friends who were also a sexual minority, (b) positively viewed and accepted their own sexuality, and (c) believed that they could choose supportive friends.

In addition, participants utilized many deception strategies to pretend to be heterosexual in front of their friends. Because society assumes heterosexuality, some LGB individuals deceived their friends by hiding their non-heterosexual orientation and letting others assume they were heterosexual (i.e., passive strategy). Others actively talked about heterosexual topics with their friends, such as making comments and jokes about persons of the opposite sex (i.e., active strategy). Also, LGB individuals directly interacted with persons of the opposite sex, via flirting, dating, and marriage, to deceive about their sexual orientation (i.e., interactive strategy). For same-sex romantic couples, individuals at times avoided introducing or denied their partners in order to hide their same-sex relationships (i.e., avoidance or denial). In contrast, some admitted their partners as important persons in their life but referred to them as "good friend," "roommate," or "cousin" (i.e., naming). In the discussion, we link several key findings to existing theories and provide suggestions for future research.

First, participants' and their friends' individual characteristics emerged as a factor that influences people's sexual orientation deception. We observed that participants who deceived their friends tended to be older, and the friends being deceived were also likely to be older. This may be because that most older adults spent much of their early adult years in a social and political environment where homosexuality was considered a mental illness and same-sex relationships were illegal (D'Augelli, Grossman, Hershberger, & O'connell, 2001). Older heterosexual adults may hold (or be perceived by the participants to have) stronger negative attitudes toward homosexuality, and older LGB adults themselves may experience greater internalized homophobia, both of which in turn could encourage sexual orientation deception. In addition, participants were less likely to deceive when their friends were also

non-heterosexual. Uncertainty reduction theory (Berger & Calabrese, 1975) contends that similarities between communicators help reduce uncertainty and promote self-disclosure. In the current context, sexual minority individuals may feel less uncertain about their LGB friends' reactions to their self-disclosure and thus are more likely to be honest.

In addition, fear of negative consequences emerged as a crucial reason that explains sexual orientation deception. Research on interpersonal conflict has indicated the central role of fear of adverse outcomes in conflict avoidance. For instance, work on the chilling effect (Cloven & Roloff, 1993; Roloff & Cloven, 1990; Solomon & Samp, 1998) shows that powerless individuals are less likely to confront their powerful partners about relational problems because underpowered individuals fear the negative responses of their over-powered partners, such as physical aggression (i.e., personal consequences) and termination of the relationship (i.e., relational consequences). The difference in sexual orientation can be viewed as a form of interpersonal conflict between friends. In this current study, LGB participants expressed concerns about whether their friends would judge them, reveal their non-heterosexual orientation to others, or terminate their friendships. Their fear of negative outcomes, thus, led them to avoid the "sexual orientation" conflict by pretending to be heterosexual.

Moreover, the findings that homophobia encouraged sexual orientation deception, whereas self-acceptance of LGB identity promoted honesty, have implications for studies on minority stress. Past research indicates that minority individuals often experience a high degree of external (e.g., friends' negative attitude to LGB people) and internal (e.g., internalized homophobia) stressors. Those stressors can give rise to various stress responses, such as anxiety and high blood pressure, that accrue over time and eventually lead to poor mental and physical health (Dohrenwend, 2000; Meyer, 2003; Pascoe & Smart Richman, 2009). Because sexual orientation deception decreases individuals' opportunities to receive social support from their friends, and social support both directly influences and buffers the negative effects of minority stress on health (Meyer, 2003), sexual orientation deception may be a communicative mechanism underlying the link between minority stressors and poor health outcomes among LGB individuals. That is, homophobia can increase sexual orientation deception, which decreases received social support from friends, eventually resulting in poor health. Future studies are encouraged to examine this prediction.

Furthermore, when it comes to sensitive topics, individuals may feel unable to communicate with others (Afifi, Olson, & Armstrong, 2005). Thus, communication efficacy has been found to influence a variety of disclosure decision-making. For instance, individuals with lower communication efficacy are less likely to disclose personal secrets (Afifi & Steuber, 2009), spouses' infertility (Steuber & Solomon, 2011), and complaints about a partner's behaviors (Worley & Samp, 2016). Consistent with these results, the data

affirmed that lack of communicative ability to talk about one's sexual orientation discouraged sexual orientation disclosure, which indirectly promoted deception. However, communication efficacy to come out is different from that to deceive. Thus, a direction for future research is to focus on LGB individuals' communicative ability to pretend to be heterosexual. For example, as there are several strategies for these deception attempts, researchers may explore whether higher communication efficacy to deceive predicts the use of more interactive strategies (e.g., flirting with a person of the opposite sex around friends) than passive tactics (e.g., keeping silent about sexual topics and letting others assume they are heterosexual).

Finally, two points are noteworthy in terms of communicative strategies identified for sexual orientation deception. First, participants seldom utilized just one strategy to deceive friends about their sexual orientation. Instead, most of them wrote two or more tactics. For example, Kobe (34, gay) stated, "Sometimes I talked about girls with them (i.e., active strategy), but most time I kept quiet when they are chatting" (i.e., passive strategy). The choices of different deception tactics may be an interesting direction for future research. For instance, goal-driven message production scholars (e.g., Berger, 1997; Dillard, 2008) contend that individuals have multiple goals during social interactions, and their choices of persuasion strategies are influenced by goal attainment. Thus, while LGB individuals have the primary goal of pretending to be heterosexual, their other secondary goals (e.g., maintaining friendships, asking for relational advice, protecting same-sex partners) may also influence which specific deception tactics they will employ. Second, we noted that even within the same strategy category, some behaviors might be more deceptive than others. For example, compared to dating or flirting with a woman, a gay man marrying a female is more likely to make his friends believe that he is heterosexual. In other words, although we coded "flirting," "dating," and "marriage" into the same category of interactive strategies, we acknowledge that they differ in degrees of deceptiveness. Thus, future studies should further investigate the differences among those deception behaviors.

This study is limited in several ways. First, we had a relatively small sample size, and the primary analyses were qualitative, limiting the generalizability of the findings. Second, we did not ask participants' relationship status. The data suggested that individuals may use different strategies to pretend to be heterosexual when they are single versus in a relationship. Thus, future studies should collect participants' romantic relationship status data and explore if certain strategies are exclusively used by single and coupled LGB individuals. Third, participants were asked to provide their responses based on the three friends they listed. This task aimed at promoting reflection on interactions with specific persons. However, friendships may differ in terms of quality and intimacy. We did not measure or ask participants' perceptions of their friendships; thus, we are unable to explore how relational factors may influence their decisions to engage in sexual orientation deception and the strategies used during those deception attempts.

CONCLUSION

The past decade has witnessed great advances in sexual minority social movements and research. However, "certain segments of American society still treat us as if we are dirt" (Kelsey, lesbian, 29), and the phenomenon of sexual orientation deception is still prevalent. This chapter serves as an initial investigation of why and how LGB individuals pretend to be heterosexual in front of their friends. More qualitative works are needed to further explore the potential advantages and disadvantages of sexual orientation deception. In addition, based upon the themes that emerged from the data, quantitative works should apply and modify existing theories to explain and predict sexual orientation deception. For instance, studies can apply the theory of planned behavior (Ajzen, 1991) and examine how internalized homophobia (an indicator of attitude to sexual orientation deception), friends' homophobia (an instance of perceived social norms), and communication efficacy and perceived negative consequence (indexes of perceived behavior control) jointly predict intentions to deceive others, which in turn is associated with deception behaviors. Moreover, future research should extend this study by focusing on other deceptive targets, such as family members, health care providers, and colleagues. Inspired by all these opportunities, we call for more examinations of the intersection between sexual minority studies and deceptive communication.

NOTES

1. We did not focus on transgender individuals because some of them (maybe due to their physical appearance or other visible markers) may not be able to conceal and/or deceive their sexual identity as their lesbian, gay, and bisexual (LGB) counterparts do. Here, "sexual minority" and LBG are used interchangeably.
2. Berger (1990) used the term "passing" to describe the "social process by which gay men and lesbians present themselves to the world as heterosexuals" (p. 328). We decided to use the term "sexual orientation deception" for two reasons. First, given that we also examined bisexual individuals' experiences of pretending to be heterosexual, sexual orientation deception seems more inclusive. Second, sexual orientation deception better aligns with the theme of this book and highlights the communicative elements in the phenomenon.
3. Recent evidence (e.g., Kees, Berry, Burton, & Sheehan, 2017; Peer, Vosgerau, & Acquisti, 2014) suggests that compared to traditional online student samples or professionally recruited panel samples, Amazon's Mechanical Turk (MTurk) samples tend to complete questionnaires with greater depth and engagement, as well as provide equal or higher-quality data. More importantly, self-reflection of sexual orientation deception may evoke discomfort, such as feelings of stress, tension, sadness, and shame. Participants may also feel being judged and criticized if reporting their deception experiences in face-to-face settings and thus prefer more private, safer ways to share their opinions and stories. Other sampling methods, such as snowball sampling and respondent-driven sampling, often require participants to be recruited by their peers. In such

cases, LGB individuals who have not come out and do not know other sex-ual minority peers may be excluded from the study. Indeed, previous studies (e.g., Papa, Lancaster, & Kahler, 2014) have documented the utility of MTurk to recruit hard-to-reach samples, including sexual minority individuals (Vaughn, Cronan, & Beavers, 2015). Thus, MTurk is an appropriate platform to recruit participants.

4. Gender and sexual orientation are often considered nominal variables. However, treating and measuring these variables as continuous is not uncommon in sexual minority studies. For instance, the Heterosexual-Homosexual Rating Scale (also known as "The Kinsey Scale"; Kinsey, Pomeroy, & Martin, 1948) measures people's sexual orientation on a 7-point scale (0 = exclusively heterosexual, 6 = exclusively homosexual). Aligning with this approach, we treated friends' gender and sexual orientation as continuous and created a hypothetical friend profile by averaging the gender and sexual orientation measures of the three real friends. This method provides for us an additional way to quantitatively explore participants' friendship characteristics.

References

Afifi, T. D., Olson, L. N., & Armstrong, C. (2005). The chilling effect and family secrets. *Human Communication Research, 31*(4), 564–598.

Afifi, T. D., & Steuber, K. (2009). The revelation risk model (RRM): Factors that predict the revelation of secrets and the strategies used to reveal them. *Communication Monographs, 76*(2), 144–176.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211.

Bagwell, C. L., Bender, S. E., Andreassi, C. L., Kinoshita, T. L., Montarello, S. A., & Muller, J. G. (2005). Friendship quality and perceived relationship changes predict psychosocial adjustment in early adulthood. *Journal of Social and Personal Relationships, 22*(2), 235–254.

Berg, N., & Lien, D. (2009). Sexual orientation and self-reported lying. *Review of Economics of the Household, 7*(1), 83–104.

Berger, C. R. (1997). *Planning strategic interaction*. Mahwah, NJ: Lawrence Erlbaum Associates.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research, 1*(2), 99–112.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage.

Berger, R. M. (1990). Passing: Impact on the quality of same-sex couple relationships. *Social Work, 35*(4), 328–332.

Berger, R. M. (1996). *Gay and gray: The older homosexual man* (2nd ed.). New York, NY: Routledge.

Braaten, L. J., & Darling, C. D. (1965). Overt and covert homosexual problems among male college students. *Genetic Psychology Monographs, 71*(2), 269–310.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Choi, S. K., & Meyer, I. H. (2016). *LGBT aging: A review of research findings, needs, and policy implications*. Los Angeles, CA: The Williams Institute.

Cloven, D. H., & Roloff, M. E. (1993). The chilling effect of aggressive potential on the expression of complaints in intimate relationships. *Communication Monographs, 60*(3), 199–219.

Cole, S. W., Kemeny, M. E., Taylor, S. E., & Visscher, B. R. (1996). Elevated physical health risk among gay men who conceal their homosexual identity. *Health Psychology, 15*(4), 243–251.

Corbin, J., & Strauss, A. (1990). Grounded theory research: Procedures, canons and evaluative criteria. *Zeitschrift für Soziologie, 19*(6), 418–427.

D'Augelli, A. R., Grossman, A., Hershberger, S., & O'connell, T. (2001). Aspects of mental health among older lesbian, gay, and bisexual adults. *Aging & Mental Health, 5*(2), 149–158.

Detrie, P. M., & Lease, S. H. (2007). The relation of social support, connectedness, and collective self-esteem to the psychological well-being of lesbian, gay, and bisexual youth. *Journal of Homosexuality, 53*(4), 173–199.

Diaz, R. M., Ayala, G., Bein, E., Henne, J., & Marin, B. V. (2001). The impact of homophobia, poverty, and racism on the mental health of gay and bisexual Latino men: Findings from 3 US cities. *American Journal of Public Health, 91*(6), 927–932.

Dillard, J. P. (2008). Goals-Plans-Action theory of message production: Making influence messages. In L. A. Baxter & D. O. Braithwaite (Eds.), *Engaging theories in interpersonal communication: Multiple perspectives* (pp. 65–76). Thousand Oaks, CA: Sage.

Dohrenwend, B. P. (2000). The role of adversity and stress in psychopathology: Some evidence and its implications for theory and research. *Journal of Health and Social Behavior, 41*(1), 1–19.

Frost, D. M., & Meyer, I. H. (2009). Internalized homophobia and relationship quality among lesbians, gay men, and bisexuals. *Journal of Counseling Psychology, 56*(1), 97–109.

Gwartney, P. A., & Schwartz, D. S. (2016). You may kiss the groom: Americans' attitudes toward same-sex marriage. In A. Machin & N. Stehr (Eds.), *Understanding inequality: Social costs and benefits* (pp. 127–149). Wiesbaden: Springer Fachmedien Wiesbaden.

Harrison, T. W. (2003). Adolescent homosexuality and concerns regarding disclosure. *Journal of School Health, 73*(3), 107–112.

Herek, G. M., Chopp, R., & Strohl, D. (2007). Sexual stigma: Putting sexual minority health issues in context. In I. H. Meyer & M. E. Northridge (Eds.), *The health of sexual minorities: Public health perspectives on lesbian, gay, bisexual and transgender populations* (pp. 171–208). New York, NY: Springer.

Herek, G. M., Gillis, J. R., & Cogan, J. C. (2009). Internalized stigma among sexual minority adults: Insights from a social psychological perspective. *Journal of Counseling Psychology, 56*(1), 32–43.

Kees, J., Berry, C., Burton, S., & Sheehan, K. (2017). An analysis of data quality: Professional panels, student subject pools, and Amazon's Mechanical Turk. *Journal of Advertising, 46*(1), 141–155.

Kertzner, R. M., Meyer, I. H., Frost, D. M., & Stirratt, M. J. (2009). Social and psychological weil-being in lesbians, gay men, and bisexuals: The effects of race,

gender, age, and sexual identity. *American Journal of Orthopsychiatry, 79*(4), 500–510.

Kinsey, A., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male*. Philadelphia, PA: Saunders.

Lyubomirsky, S., Tkach, C., & DiMatteo, M. R. (2006). What are the differences between happiness and self-esteem. *Social Indicators Research, 78*(3), 363–404.

McLean, K. (2001). Living life in the double closet: Bisexual youth speak out. *Hecate, 27*(1), 109–118.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129*(5), 674–697.

Morris, J. F., Waldo, C. R., & Rothblum, E. D. (2001). A model of predictors and outcomes of outness among lesbian and bisexual women. *American Journal of Orthopsychiatry, 71*(1), 61–71.

Myrick, F. (1974). Attitudinal differences between heterosexually and homosexually oriented males and between covert and overt male homosexuals. *Journal of Abnormal Psychology, 83*(1), 81–86.

Nardi, P. M. (1999). *Gay men's friendships: Invincible communities*. Chicago, IL: University of Chicago Press.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Papa, A., Lancaster, N. G., & Kahler, J. (2014). Commonalities in grief responding across bereavement and non-bereavement losses. *Journal of Affective Disorders, 161*(1), 136–143.

Pascoe, E. A., & Smart Richman, L. (2009). Perceived discrimination and health: A meta-analytic review. *Psychological Bulletin, 135*(4), 531–554.

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46*(4), 1023–1031.

Peplau, L. A., & Fingerhut, A. W. (2007). The close relationships of lesbians and gay men. *Annual Review of Psychology, 58*, 405–424.

Peplau, L. A., & Spalding, L. R. (2000). The close relationships of lesbians, gay men and bisexuals. In C. Hendrick & S. S. Hendrick (Eds.), *Close relationships: A sourcebook* (pp. 111–124). Thousand Oaks, CA: Sage.

Roloff, M. E., & Cloven, D. H. (1990). The chilling effect in interpersonal relationships: The reluctance to speak one's mind. In D. Cahn (Ed.), *Intimates in conflict: A communication perspective* (pp. 49–76). Hillsdale, NJ: Lawrence Erlbaum.

Rosario, M., Hunter, J., Maguen, S., Gwadz, M., & Smith, R. (2001). The coming-out process and its adaptational and health-related associations among gay, lesbian, and bisexual youths: Stipulation and exploration of a model. *American Journal of Community Psychology, 29*(1), 133–160.

Solomon, D. H., & Samp, J. A. (1998). Power and problem appraisal: Perceptual foundations of the chilling effect in dating relationships. *Journal of Social and Personal Relationships, 15*(2), 191–209.

Stanley, J. (1996). The lesbian's experience of friendship. In J. S. Weinstock & E. D. Rothblum (Eds.), *Lesbian friendships* (pp. 39–59). New York, NY: New York University Press.

Steuber, K. R., & Solomon, D. H. (2011). Factors that predict married partners' disclosures about infertility to social network members. *Journal of Applied Communication Research, 39*(3), 250–270.

Ueno, K. (2005). Sexual orientation and psychological distress in adolescence: Examining interpersonal stressors and social support processes. *Social Psychology Quarterly, 68*(3), 258–277.

Vaughn, A. A., Cronan, S. B., & Beavers, A. J. (2015). Resource effects on in-group boundary formation with regard to sexual identity. *Social Psychological and Personality Science, 6*(3), 292–299.

Wells, J. W., & Kline, W. B. (1987). Self-disclosure of homosexual orientation. *The Journal of Social Psychology, 127*(2), 191–197.

Worley, T. R., & Samp, J. (2016). Complaint avoidance and complaint-related appraisals in close relationships: A dyadic power theory perspective. *Communication Research, 43*(3), 391–413.

Student Cheating: A Dramaturgical Analysis of Identity, Deception, and Self-deception

Susan A. Stearns

Academic dishonesty is a pervasive problem in higher education. Approximately 65–100% of students admit to cheating at least once during college (Stearns, 2001, p. 275). The vast majority of studies examining student cheating are quantitatively oriented, offering both insights and various shortcomings. Scholars seek to understand factors contributing to academic dishonesty (Giluk & Postlethwaite, 2015; McCabe, Feghali, & Abdallah, 2008), as well as how to create a culture of academic integrity (Mansoor & Ameen, 2016; VanDeGrift, Dillon & Camp, 2017). Universities and researchers have also become concerned with online academic dishonesty (e.g., Sendag, Duran, & Fraser, 2012).[1]

The current study offers an alternative, fresh approach to the literature on academic dishonesty. In this study, 72 students responded to an online qualitative questionnaire on academic integrity. Although each participant acknowledged having *never* cheated while in college prior to being granted access to the survey, seven participants later reported cheating. Using Goffman's dramaturgical perspective, I offer a discourse analysis of these seven participants' responses. Each participant's response is examined as a deceptive performance. Taken together, the participants' performances can be understood as creating three distinct social identities. These identities offer opportunities to further investigate the complexities of deception. These complexities involve: (1) the depth of deception (i.e., from entering the study by claiming to have never cheated to increasingly deceptive behaviors

S. A. Stearns (✉)
Eastern Washington University, Cheney, WA, USA
e-mail: sstearns@ewu.edu

exhibited in responses to questions in the study), (2) the perception of who the audiences are for the dramaturgical performance, and (3) considering the role of self-deception when performing for the self as audience. First, I offer an overview of the dramaturgical perspective, which is used later to analyze students' discourse.

The Dramaturgical Perspective

Dramaturgy, as presented by Goffman (1959), depicts individuals as performers requesting audience members to accept their performance as legitimate (i.e., to be believed as reality, p. 17). The front stage includes the performer's time in front of the audience with the potential to influence audience members (p. 22). In traditional dramaturgy, the front stage includes *setting*, *appearance*, and *manner*. For this analysis, the reliance is upon *manner* since the setting and appearance are only minimally apparent in an online survey format. In this discourse analysis, word choice, tone, and the respondent's consistency between their own responses (Goffman, p. 24) assist in the investigation of deception. This study will also rely heavily upon Scott and Lyman's (1968) work on *accounts* which is an extension of Goffman's dramaturgical theory. An account is "…a statement made by a social actor to explain…untoward behavior" (Scott & Lyman, p. 46). The cheating behaviors reported by the respondents are the untoward cheating behaviors that will be analyzed in this study.

A number of additional dramaturgical concepts are important in this analysis—the study's non-interactional setting, a performance as a dress rehearsal for future performances, and the self as audience. The relevance of each of these three concepts is discussed next. Goffman (1959) thoroughly addresses the precarious position performers place themselves in within an interactional setting: "…at any moment in their performance an event may occur to catch them out and baldly contradict what they have openly avowed, bringing them immediate humiliation and sometimes permanent loss of reputation" (p. 59). This is because of, as Conte (2008) notes, the "…dynamic relationship of socially situated reciprocal actions" (p. 379). Since the dynamic relationship offered in a typical interaction described by Goffman is nonexistent in this non-interactional online survey setting, the performer no longer faces a precarious situation. The distant audience has no ability to call the deceiver to account. This type of audience offers a freedom for the performer—whatever the performer asserts cannot be challenged.

The fact that the audience is distant and incapable of interacting with the performer can alter the perspective of the performer in two distinct ways. First, the individual may view his/her performance as a dress rehearsal for future interactive performances. Second, this necessitates the individual becoming his/her own audience. In regard to the dress rehearsal concept, it may be that the responses in this study are more than a performance: Could

the participants' responses be a rehearsal for future perfected deceptive performances (Goffman, 1959, p. 156)? This question arises because while each performer's discourse can be analyzed, the rationale for why the performer is participating in this study is under question. As Goffman describes rehearsals, "…performers can become practiced in their parts and so that contingencies that were not predicted will occur under circumstances in which they can be safely attended to" (p. 228). Thus, a rehearsal is a safe place for a performer to practice his/her art. The non-interactive setting of an online survey offers the performer a safe place where he or she can become "practiced" in their utterances and proclamations of innocence without risk of being challenged and potentially humiliated.

Secondly, as noted above, the individual needs to become their own audience for this dress rehearsal. The performer needs to know how well he or she has controlled the information conveyed (Shulman, 2017, p. 229) and to what extent the performance is believed. In an online survey, only the performer's self can judiciously evaluate the performance since the audience, in this case the researcher, is at a distance. This distance disallows feedback. The audience of the self "attends" to his/her own responses and evaluates the potential for effective deception. The self advises how well the performer has controlled his/her information.

Greenwald and Breckler (1985) argue the *primary* audience for performances is the self (i.e., the inner-audience hypothesis, p. 126). This chapter argues not that the self is the primary audience at all times, but rather that the primary audience fluctuates in a performer's mind based on the context, topic, and situation. One's performance can be directed at one audience or multiple audiences simultaneously. One of the arguments presented within this chapter is that when the performance may in reality be a combined performance and rehearsal for further performances, the self as audience becomes a necessity in order to obtain feedback. It is this feedback that will encourage the performer to alter his/her future behaviors to present a more believable future performance.

A Discourse Analysis of Apparent Deception

The original study (Stearns, 2007) requested participants who fulfilled the following criteria: (1) eighteen years of age or older; (2) currently enrolled in college or had been enrolled in college during the last year; and (3) had never cheated on academic work while in college. Participants agreed to all three criteria to access the online survey. The survey itself consisted of fifteen open-ended questions that focused on how their choice not to cheat affected their relationships with peers, parents, and faculty.[2] Due to Institutional Review Board concerns about the sensitive nature of issues regarding academic integrity and dishonesty, demographic data pertaining to these students were not collected.[3]

Typical of discourse analysis, the responses from each participant were carefully analyzed to determine common themes (Wood & Kroger, 2000). Each of the seven respondents either acknowledged or alluded to their lack of academic integrity. Consistent with Presser's (2004) findings, distinct identities emerged from the respondents' accounts of their deceptive behaviors. Three identities emerged: *the entertainer*, *the confessor*, and *the justifier*. Each social identity is discussed below.

The Entertainer

The socially constructed identity of the entertainer arose from the accounts of only one respondent, Pat. To contextualize Pat's entire set of remarks, understanding Pat's response to the final survey question is essential. When asked, do you have anything else to share? Pat remarked, "lulz." According to the *Oxford Dictionary*, "lulz" refers to "Fun, laughter, or amusement, especially that derived at another's expense." Thus, it appears that Pat's accounts are offered to me, the researcher, for self-entertainment and intended to be at the researcher's expense.

Interpreting Pat's remarks within the understood context of "lulz," note the repetitive use of sarcastic humor in the following responses. When asked how Pat defined cheating, Pat stated, "With a dictionary." When asked why Pat had consistently chosen honesty, Pat responded, "I've been honest?" When asked how the choice of honesty affected relationships with professors (and also parents), Pat responded to both questions, "They love me for having straight A's." And when asked what prompts Pat to speak to peers about not cheating, Pat responded, "I don't. I encourage them to cheat." The tone of each of these remarks is humorous with an interwoven derision toward the research theme and/or the researcher. Goffman (1959) refers to these types of utterances as *aggressions* (pp. 174–175).

In the next set of responses, Pat's performance extends beyond sarcastic humor to what Goffman (1959) calls *derisive collusion*. Derisive collusion typically occurs between an individual and his/her self (e.g., the child tells a lie while crossing their fingers behind their back or the employee who grimaces at the boss when the boss turns around, p. 187). In a way, Pat taunts the researcher by mocking academic integrity and admitting to cheating. When asked, what has helped you remain honest when tempted to cheat, Pat responded, "I cheated." When asked whom do you talk to about your decision not to cheat, Pat emphatically responded, "I DO cheat." And when asked, what is the closest you've come to cheating, Pat responded, "Actually cheating." This mockery of sorts emphasizes Pat's derisive performance. But the distant nature of this online survey allows Pat to hide behind anonymity. Pat taunts and mocks the distant audience, the researcher, from afar.

The entertainer's presentation of self confesses his/her deception: Pat cheated while in college. Pat's overall presentation is one of self-satisfied

mockery and sarcastic humor; Pat acts as an entertaining court jester. Avner and Gorsky (2006) note, "…online deception seems to be an enjoyable activity" (p. 58). They add, "This may be associated with the medium: It is certainly less threatening to deceive someone you don't know and, if you so desire, will never know" (p. 58). Pat discovered a venue for putting on this performance that Goffman did not imagine; this stage allows no one to baldly contradict Pat's apparently enjoyable performance.

Unfortunately, the enjoyment of deception is not a typical topic in the deception literature (Van Dongen, 2002, p. 149). Yet there are individuals, like Pat, who take great delight in duping others whether in face-to-face or online settings. Ford (1996) discusses these individuals:

> The delight that one may achieve through fooling someone else is one aspect of achieving a sense of power through deceit…A thrill is often associated with fooling someone else – a sense of power, cleverness, and superiority. Ekman (1992) coined the phrase *duping delight* to describe this particular form of pleasure in successfully perpetrating a deceit. Bursten (1972) described similar feelings by his use of the term *putting one over.* (p. 92; italics added)

McEntire (2002) also notes this phenomenon in his multi-societal study of tricksters:

> The first requirement is that the prankster will succeed in bringing one, several, or many people into an unreal world–for a few seconds, for a moment, or longer, and that the deception not cause needless harm. The second requirement is that the deception will be revealed to the victim or victims. (p. 145)

This repeated pattern McEntire refers to occurs with the entertainer. First, Pat invites the audience into this unreal world. Pat then fulfills the second criterion by interweaving the "reveal" throughout the utterances culminating with, "lulz."

Social scientists rarely address the sheer joy or benefits of deception. An extraordinary amount of scholarly work examining deception focuses on deception detection, as well as the role of cognitive load (i.e., that lying requires extra mental effort).[4] Yet the figure of the entertainer brings to light a clear dearth in the deception literature—an understanding of everyday interactions where people choose to deceive for enjoyment. Indeed, one reason people lie is to have fun (Knapp, 2008), and the purpose of deception is, at times, for entertainment. However, social science has yet to explore these ideas in much depth.

The Confessor

The dataset comprised two confessors—Jamie and Aubrey—and they represent themselves in distinctly different manners. In the following accounts,

note Aubrey's direct and apparently honest responses. When asked, "Why do you think you've been consistently honest in college?" Aubrey replied, "I have cheated." Furthermore, Aubrey readily admits to cheating in *all* responses. But Aubrey offers more than mere confession; Aubrey provides an account pertaining to when it is appropriate to cheat. First, Aubrey articulates a personally acceptable level of cheating as the threshold to determine if another person's behavior is acceptable or not:

> Well if they cheat every now and then I don't care but if they cheat on everything which I've seen before I looked down on them. It's pathetic if you cheat on everything. You have to learn to do things on your own sometimes. I mean yeah I've cheated but not on everything. I once had a friend who would cheat on everything. Copy homework, copy others on test, plagerize on papers, she did this on everything. I started to lose respect for her.

Aubrey's personal theory begins with the idea that some cheating is acceptable. But once Aubrey designates someone's type of cheating as "on everything" (i.e., beyond Aubrey's personal threshold), Aubrey has issues with said person. This was exhibited in another response as well: "…unless they cheat all the time then I just look down on that." Aubrey differentiates from others who cheat more: "…plus I don't want to be a dumbass who cheats all the time and doesn't know shit about anything." Aubrey's full confession embedded within this theory of academic integrity attempts to mitigate the audience's perception of his/her behavior and, in doing so, explains the rationale behind the deception.

Jamie, on the other hand, is an *unexpected confessor*. Jamie responded briefly to the first thirteen questions with no reference to having cheated, but when asked about the closest Jamie had ever come to cheating Jamie simply responded, "cheated." One small word, not even capitalized. Then, as if this confession opened a floodgate of emotion, in the very next and last response, which asked if there was anything else to share, Jamie stated, "Tell someone who cheated that if they cheat again someone will punch them in the fucking teeth." This utterance's seemingly metaphorical demand for acute physical violence is filled with a volatile anger not present in Jamie's previous responses.

Jamie's confession appeared with no further communicative offerings as to why the performance abruptly altered from an honest student to a confessed cheater. According to Peer, Acquisti, and Shalvi (2014), Jamie falls within the realm of being a *partial confessor*, which they define as when "…people restrict their honesty about their prior dishonesty" (p. 202). Unfortunately for partial confessors, "…partially confessing actually aggravates, rather than alleviates, negative feelings" (Peer et al., p. 213). These negative feelings may help to explain Jamie's violent outburst in the response after confessing.

Aubrey's response, by contrast, was a full confession. Horowitz (1956) observes, "…'Confession is good for the soul.' Apparently, cathartic and

purgative psychological properties accrue from confession" (p. 197). While Aubrey's confessional performance of confiding to the researcher may have resulted in a cathartic release, Jamie's partial confession may have aggravated an emotional state. Both confessors appear to attempt to fulfill a compulsion to disclose their cheating behaviors on a safe stage where there were no consequences.

To date, the deception literature examining confessions focuses primarily in legal contexts—false confessions, forced confessions, and confessions when confronted with evidence demonstrating deceit during an interrogation (Kassin & Gudjonsson, 2004; Shuy, 1998). Yang, Guyll, and Madon (2017) offer a rationale for this honed focus on legal confessions: "In the criminal justice system, a confession is among the most persuasive forms of incriminating evidence" (p. 80). The serious consequences of a legal confession have demanded research attention and yet share very little commonality with everyday performances of confessions such as non-requested confessions and unexpected confessions. Studying the everyday occurrence of confessions, rather than the legal aspects of confession, could substantially add to our knowledge of deception—why do people deceive and then suddenly disclose their deception? What are the reasons for unprompted versus prompted confessions, and how, if at all, does discourse differ in these situations? What happens when a person has the opportunity to rehearse his or her confession, as compared to having no opportunity to rehearse?[5] These types of questions are largely unaddressed in the extant deception literature.

The Justifier

Four respondents, labeled *justifiers*, offered a variety of defenses for their cheating behavior. Lee, for example, referred to a take-home test where students were required to work individually:

> Like I said, that take-home test would be counted by many as cheating. But we were actually learning Chemistry together, and we spent many hours poring over the text and the math of it, trying to get it nailed down, engaged in active academic discussion. So I don't count it, but you might and throw my survey out. Other than that, I'm not sure I've come close.

Lee attempts to make this account sound reasonable, even to the point of declaring that the reported behavior should not be categorized as cheating. Yet, Lee defines cheating as, "Disobeying the rules, written or understood." Lee's own definition of cheating contradicts his/her attempt to justify this behavior as acceptable. In a similar vein, another participant, Casey, wrote about the "closest" he/she ever came to cheating: "hrmm, well, I did have to write a 20 page paper for one course, and I just elaborated a lot (12 pages worth) on my term paper from high school. It was still my work but it wasn't 'fresh' research." Goffman (1959) argues, "the performer can rely upon his

audience to accept minor cues as a sign of something important about his performance" (p. 51). The use of "hrmm" is that minor cue. Casey flagged the audience as to the unusual nature of this response by creating a grammatically unknown qualifier to signal the importance of the upcoming justification account: "it was still my work." Another participant, Kelly, stated, "i copyd someones math home work perhaps but math was soooooooooooooo long ago that i can't even remmeber i'm graduate student now. one who obviously can't spell or type well. Lol." Kelly has a number of justifications supporting this argument that the behavior reported is not cheating: (1) the claim of no memory; (2) an attempt to create distance from the event; and (3) finally the offer of humor to take attention away from this convoluted claim.

The commonality in Lee, Casey, and Kelly's responses revolves around the concept of *justification*. Scott and Lyman (1968) define justifications as "accounts in which one accepts responsibility for the act in question, but denies the pejorative quality associated with it" (p. 47). Each respondent acknowledged a specific cheating event and then attempted to lessen the audience's negative view of the event. Additionally, the rationales offered for each justification fulfill Antaki's (1994) requirement for justifications: that the justifier offers a permissible reason for his or her offense (p. 47).

The responses of Morgan, also identified as a justifier, indicate a desire to obtain the esteem perceived to reside with those who chose honesty while in college. Early in the responses, Morgan admits to cheating: "I have been passed notes and answers before." Then, when asked if there was a point in life when a decision was made not to cheat, Morgan said, "well, i used to cheat and steal a lot when i was very young, about 9 years old. I just kinda woke up one morning and i stopped. Also a few years ago, i did cheat once and i was caught. I felt horrible, i talked with the teacher why it happened and he understood, but i never forgot. I haven't since then." Morgan clearly admits to *repeated* instances of cheating while in college, providing a reference to being passed notes and answers, and a disclosure about another instance of cheating that resulted in being caught. As with the other justifiers, Morgan offers a permissible reason for these cheating events when describing their emotional toll. Thus, Morgan's performance has fulfilled both the definitional and permissible reasons requirement of Scott and Lyman (1968) and Antaki (1994).

The larger deception literature examines justification in a much broader sense than discussed here. Justification has been examined in research on communication (Van Swol & Braun, 2014a, 2014b), deception detection (Paik & Van Swol, 2017; Van Swol & Braun, 2014a), and academic integrity (O'Rourke et al., 2010), to name a few. Yet, these analyses of deception and justification introduce the concept of justification devoid of the substantive work of Scott and Lyman (1968). This is particularly intriguing because Scott and Lyman's work offers a categorization of justification types, a contrast

between justification and excuses, and an explanation of when and why people use justifications. This work on justification is particularly useful for application to the topic of deception.

Deception Performances

The above discourse analysis details three distinct social identity performances: the entertainer, the confessor, and the justifier. Each of these socially constructed identities offers further understanding of how deception can be performed. Additional relevant deception concepts emerged from these social identities: audiences and self-deception.

Audience

The dramaturgical perspective recognizes the centrality of the audience to deception; deception is conceived for a particular audience or audiences and often acted out to that audience or audiences. This study recognizes three potential audiences: (1) the researcher, (2) the imagined honest students (what Goffman calls the unseen reference group), and (3) the self. Note that the consent form for this study designated the researcher as the original audience and all respondents directed their remarks toward the researcher either directly or by implication. Lee directly speaks to the researcher stating, "…So I don't count it, but you might and throw my survey out…" Lee is concerned the researcher will identify the behaviors previously described in the survey as cheating and remove all of Lee's responses from the study. The use of the word "you" exemplifies one of the few comments made directly to the researcher, thereby invoking the researcher as audience. The vast majority of the respondents spoke to the researcher by implication, not using personal pronouns.

The second type of audience resides in the performer's mind. Goffman (1959) remarks:

> When a performer guides his private activity in accordance with incorporated moral standards, he may associate these standards with a *reference group* of some kind so that, in a sense, there will be a non-present audience for his activity. This possibility leads us to consider a further one. The individual may privately maintain standards of behavior which he does not personally believe in, maintaining these standards because of a lively belief that an unseen audience is present which will punish deviations from these standards. In other words, an individual may be his own audience or may imagine an audience to be present. (pp. 81–82; italics added)

The reference group audience in the current study consists of the individuals invited to participate, the college students who had not cheated, as contrasted with the uninvited respondents, the cheaters. Morgan's justification

for deception revolves around the desire to be a member of the invited audience, the honest students' reference group. Morgan's justifications are addressed to the researcher, the honest students, and also to Morgan's self. Morgan seeks confirmation that the behaviors are honest (e.g., when stating the researcher may throw the survey out due to the reported behaviors). Morgan seeks acceptance into the invited honest student audience by responding to a survey for which Morgan did not qualify (Morgan admitted to cheating).

A third audience is the self. The "lulz" response of the entertainer, Pat, situated the self as an audience; the entertainer aimed to provide entertainment for the self and perhaps also the researcher. Any performer conducting a rehearsal also includes the self as an audience. And finally, the justifiers also focused on the self as an audience, but in a slightly different manner. When a justifier attempts to offer permissible reasons for their behavior, the self then must "test" whether the permissible reason is acceptable. The self must also test the acceptability of the overall performance, especially when there is a non-interactive audience.

A performer will attend to the audiences they perceive; for some performers, this means a singular audience, whereas other performers may perceive multiple audiences. A performer's perception of who the audience is can and will change any performance, especially a deceptive performance since a poorly designed account could damage their identity of self. As Goffman (1959) expresses, "We will also find that the circumspect performer will attempt to select the kind of audience that will give a minimum of trouble in terms of the show the performer wants to put on and the show he does not want to have to put on" (pp. 218–219).

Social scientists will benefit from examining who the audiences might be for any individual performance, since performers craft their communication to their perceived audiences. A fuller understanding of the communication analyzed in deception studies is possible when we begin to see that the performance is directed at an audience or audiences that are physically present, as depicted in typical face-to-face interactions, a non-present reference group, and/or the performer's self.

Self-deception

The importance of who the individual perceives as the audience(s) offers the potential to more fully contemplate self-deception. The specific focus herein is upon deceptive performances that include the self as audience, which I argue increases the opportunity for self-deception. A discourse analysis begins this discussion.

When asked, "How do you think your decision not to cheat has affected your relationship with your parents?" each justifier below responded as if they were honest and had never cheated while in college:

Lee: It's more the other way around. Momma taught me not to cheat, so I don't. Because I wouldn't want her to know I'm doing that, and I would even less like to lie to her. I've lied to the dear woman before, and it's about my least favorite thing in the world.

Kelly: my parents have always trusted me and cheating is like lying if you want to get ahead in life you don't do it. you are honest and follow the rules even if they are stupid.

Casey: Again, I don't think my parents would ever think I would. My mom would definitely be "verrrry disappointed" in me if I were caught cheating.

The accounts above explain why each participant chose to be honest; when analyzing the discourse within an individual response, the performance appears legitimate. Yet, when analyzing these participants' discourse across responses to the survey questions, as done earlier in this chapter, Lee, Kelly, and Casey all admitted to at least one cheating incident while in college.

One of the paradoxes of self-deception occurred as each performer presented accounts of honesty, while offering detailed accounts of their cheating behavior in other responses. Since Lee, Kelly, and Casey are all justifiers and one of the audiences of the justifiers is the self, I propose each of these justifiers attempted to deceive themselves. Goffman (1959), in reference to when the performer comes to be his or her own audience, explained:

> In these cases it will have been necessary for the individual in his performing capacity to conceal from himself in his audience capacity the discreditable facts that he has had to learn about the performance; in everyday terms, there will be things he knows or has known, that he will not be able to tell himself. (pp. 80–81)

The audience centeredness of dramaturgy strongly encourages the audience to be restricted from viewing backstage behavior. Minor discrepancies in a performance can destroy the audience's belief in the performance. Yet, the performer who is also an audience (i.e., the self audience) is thrust into the position of being both front stage and backstage. This predicament leads to the performer confronting discreditable facts about the self and attempting to believe or coming to a belief in their own deceitful performance if they are to protect their self-identity.

Furthermore, the opportunity for a dress rehearsal may assist the self in denying cheating behaviors and embracing the identity of an honest college student. For as Blumstein et al. (1974) state, when an offender puts forward an unchallenged account, "…it is tantamount to his accepting the identity into which he has been cast. Not only are identities established and modified in this way, but the remedial process often reflects and reaffirms identities already established" (p. 553). In the 1970s, Blumstein obviously did not imagine online surveys like the one used for the current study, but he did envision a situation wherein no one would challenge an individual's

performance, thereby encouraging the individual to embrace their proposed identity. The very act of creating and performing identities wherein the accounts cannot be challenged encourages the individual to potentially partake in self-deception, accepting their own deceit as plausible.

The broader literature on deception further delves into the complexities of self-deception (Solomon, 1993):

> Deception and self-deception, I want to argue, are conceptually distinct, but thoroughly entangled phenomena…To fool ourselves, we must either fool or exclude others; and to successfully fool others we best fool ourselves…. Transparency to ourselves can be just as intolerable as transparency to others and for just the same reason…Part of the self is self-presentation and self-disclosure, but an aspect of equal importance is the need to hide, not to disclose, those facets of the self that are less than flattering, humiliating, or simply irrelevant to the social context or interpersonal project at hand. (p. 42)

The need to hide aspects of the self is how we create desired social identities; we self-deceive to create an identity we are proud to claim. Notice how this quote on self-deception aligns with a dramaturgical view: One can readily see the self as the audience for self-deception and when we "fool others" as a reference to another audience. Baumeister (1993) reiterates these points when he notes:

> A person who wants to believe something finds it easier to believe if others can be induced to believe it, or even just to go along with it. Lying to others can thus be a means of lying to oneself. On the other hand, a carefully woven network of self-deceptions can be rudely disrupted if other people fail to validate it or point out its fallacies. Self-deception is thus easiest to accomplish in the privacy of one's own mind. But if other people become involved it becomes urgently necessary to convince them as well. (Baumeister, 1993, p. 177)

Baumeister adds one additional argument when he notes this quote is in regard to *people who want to believe something*. This desire to believe may assist the performer in not just attempting to believe their deception, but in actually coming to believe their deception. Harking back to an earlier section of this chapter, remember Morgan's desire to be an honest college student; Morgan really wanted to be the individual who changed and behaved honestly since nine years of age. Yet Morgan's responses were a pretense of honesty riddled with references to cheating.

Vital to our understanding of dramaturgy's view of self-deception and the broader literature on self-deception, as represented by Solomon (1993) and Baumeister (1993), is the ability to intertwine the knowledge learned from both. Unfortunately, the deception literature appears not to utilize the dramaturgical literature, which could offer various benefits. For instance, Chance, Norton, Gino, and Ariely's (2011) study delves into the short-term and

long-term costs and benefits of self-deception. The authors repeatedly frame the behaviors of the participants as a form of performance: How they perform on individual tests to their expected performance on future tests. Yet, there is no acknowledgment of the simultaneously deeper and broader usage of "performance" in the dramaturgical literature. Additionally, Chance et al.'s analysis of intrapersonal self-deception compared to interpersonal self-deception parallels dramaturgy's acknowledgement of multiple audiences including the self. Chance et al., and possibly other scholars studying self-deception, might very well benefit from utilizing that literature on dramaturgy.

Conclusion

I agree with Scheff (2006) when he says, "…Goffman's writing shatters the calm surface of everyday life, it notices and comments upon what it is to be taken for granted by members in good standing" (p. 26). The dramaturgical analysis presented in this chapter provokes intellectual curiosity regarding deception by redirecting our attention to overlooked, everyday enactments of deception. The respondents in the current study voluntarily chose to deceive with no urging from others, nor in an attempt to protect themselves.

The analysis of actual deception performances in the current study resulted in three social identities—the entertainer, the confessor, and the justifier. Each of these identities offers distinctly different representations of deception. Studying the "taken for granted" nature of everyday deception opens our minds to additional possibilities for research in both breadth and depth. The entertainer brings to mind what earlier researchers have labeled as duping delight (e.g., Ford, 1996). This form of deception is worthy of further consideration by researchers. For example, *catfishing*, a type of deceptive activity where a person creates a fake identity on a social network account—often for vicious or nefarious purposes—may be considered a form of duping delight in the digital era. If we work even further to extend our research to deception in everyday life, we may come to better understand phenomenon such as this and prevent the success of swindlers, for example.

The confessors in the current study, for example, brought our attention to the need for individuals to purge themselves from guilt. What more could we discover about confessions, especially in various contexts, if we were to examine when and how they occur in everyday life? The current study also calls to our attention that there is much more to understand about justification—when, why, and how it is used depending on the communication context?

Additionally and possibly the most important aspect of the dramaturgical perspective in relation to deception is the focus on audiences. Specific to this study are the self as audience, the non-present reference audience, and the researcher as audience. The audience is a central element of communicative interactions, providing rich contextual information. Recognizing that

performers may cognitively acknowledge present and non-present audiences is essential to our interpretation of their deceptive performances. Understanding who the performer views as their audience(s) can only enlighten our understanding of how individuals choose to deceive. Related, when the self is recognized as a viable audience, in conjunction with understanding an individual's desire to maintain a positive self-identity, examining the complexities of self-deception becomes more fruitful.

Goffman does shatter our assumptions about everyday life as Scheff (2006) rightly observes. The mundane is worthy of study. Social scientists will benefit from following his example, especially as they study the complex performance of deceptive behavior and thought.

Notes

1. An extensive bibliography of over 1800 academic resources is housed at the International Center for Academic Integrity (ICAI). This author, Stearns, created this resource in 2000, maintained it for public access by adding resources until 2006, and now it is maintained by, and accessible to, members of the ICAI (www.academicintegrity.org).
2. This study mirrored an earlier study wherein repetitive cheaters, rather than non-cheaters, were studied (Stearns & Cantu, 2006).
3. Due to IRB constraints, gender is unknown in this study.
4. See Levine and McCornack (2014, p. 436) for a well-articulated argument against the acceptance of cognitive load research.
5. This last question focusing on rehearsals is addressed within the deception literature on the criminal justice system (e.g., Gawrylowicz et al., 2016), but not within the broader literature examining deceptive situations in everyday life.

References

Antaki, C. (1994). *Explaining and arguing: The social organization of accounts*. Thousand Oaks, CA: Sage.

Avner, C., & Gorsky, P. (2006). Online deception: Prevalence, motivation, and emotion. *Cyber Psychology & Behavior, 9*(1), 54–59.

Baumeister, R. F. (1993). Lying to yourself: The enigma of self-deception. In C. Saarni & M. Lewis's (Eds.), *Lying and deception in everyday life* (pp. 166–183). New York: Guildford Press.

Blumstein, P. W., Carssow, K. G., Hall, J., Hawkins, B., Hoffman, R., Ishem, E., …, Zimmerman, D. L. (1974). The honoring of accounts. *American Sociological Review, 39*(4), 551–566.

Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences of the United States of America, 108*(3), 15655–15659.

Conte, M. (2008). Little naked pangs of the self: The real performance of the self and the function of trust in Goffman's action theory. *International Review of Sociology, 18*(3), 375–392.

Ford, C. V. (1996). *Lies! Lies!! Lies!!! The psychology of deceit*. Washington, DC: American Psychiatric Press.

Gawrylowicz, J., Fairlamb, S., Tantot, E., Qureshi, Z., Redha, A., & Ridley, A. M. (2016). Does practice make the perfect liar? The effect of rehearsal and increased cognitive load on cues to deception. *Applied Cognitive Psychology, 30,* 250–259.

Giluk, T. L., & Postlethwaite, B. E. (2015). Big five personality and academic dishonesty: A meta-analytic review. *Personality and Individual Differences, 72,* 59–67.

Goffman, E. (1959). *Presentation of self in everyday life*. New York: Anchor Books.

Greenwald, A. G., & Breckler, S. J. (1985). To whom is the self presented? In B. R. Schlenker's (Ed.), *The self and social life* (pp. 126–145). New York: McGraw-Hill.

Horowitz, M. W. (1956). The psychology of confession. *Journal of Criminal Law and Criminology, 47*(2), 197–204.

Knapp, M. L. (2008). *Lying and deception in human interaction*. Boston, MA: Pearson Education.

Kassin, S. M., & Gudjonnsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest, 5*(2), 33–67.

Levine, T. R., & McCornack, S. A. (2014). Theorizing about deception. *Journal of Language and Social Psychology, 33*(4), 431–440.

Lulz. (2018). *Oxford Dictionary*. Retrieved January 28, 2018, from https://en.oxforddictionaries.com/definition/lulz.

Mansoor, F., & Ameen, K. (2016). Promoting academic integrity in South Asian research culture: The case of Pakistani academic institutions. *South Asian Studies, 31*(2), 77–90.

McCabe, D. L., Feghali, T., & Abdallah, H. (2008). Academic dishonesty in the Middle East: Individual and contextual factors. *Research in Higher Education, 49,* 451–467.

McEntire, N. C. (2002). Purposeful deceptions of the April Fool. *Western Folklore, 61*(2), 133–151.

O'Rourke, J., Barnes, J., Deaton, A., Fulks, K., Ryan, K., & Rettinger, D. A. (2010). Imitation is the sincerest form of cheating: The influence of direct knowledge and attitudes on academic dishonesty. *Journal of Ethics and Behavior, 20*(1), 47–64.

Paik, J. E., & Van Swol, L. M. (2017). Justifications and questions in detecting deception. *Group Decision and Negotiation, 26,* 1041–1060.

Peer, E., Acquisti, A., & Shalvi, S. (2014). "I cheated, but only a little": Partial confessions to unethical behavior. *Journal of Personality and Social Psychology, 106*(2), 202–217.

Presser, L. (2004). Violent offenders, moral selves: Constructing identities and accounts in the research interview. *Social Problems, 51*(1), 82–101.

Scheff, T. J. (2006). *Goffman unbound! A new paradigm for social science*. Boulder: Paradigm Publishers.

Scott, M. B., & Lyman, S. M. (1968). Accounts. *American Sociological Review, 33*(1), 46–62.

Sendag, S., Duran, M., & Fraser, M. R. (2012). Surveying the extent of involvement in online academic dishonesty (e-dishonesty) related practices among university students and the rationale students provide: One university's experience. *Computers in Human Behavior, 3,* 849–860.

Shulman, D. (2017). *The presentation of self in contemporary social life.* Thousand Oaks, CA: Sage.

Shuy, R. W. (1998). *The language of confession, interrogation, and deception.* Thousand Oaks, CA: Sage.

Solomon, R. C. (1993). What a tangled web: Deception and self-deception in philosophy. In C. Saarni & M. Lewis's (Eds.), *Lying and deception in everyday life* (pp. 30–58). New York: Guildford Press.

Stearns, S. A. (2001). The student-instructor relationship's effect on academic integrity. *Journal of Behavior & Ethics, 11,* 275–286.

Stearns, S. A. (2007). *Not ever: Eschewing academic dishonesty.* Paper Presented at the Meeting of the Center for Academic Integrity, Newport News, VA.

Stearns, S. A., & Cantu, J. (2006). *More than once: Repetitive acts of academic dishonesty.* Paper Presented at the Meeting of the Center for Academic Integrity, Boulder, CO.

VanDeGrift, T., Dillon, H., & Camp, L. (2017). Changing the engineering student culture with respect toward academic integrity and ethics. *Science and Engineering Ethics, 4,* 1159–1182.

Van Dongen, E. (2002). Theatres of the lie: 'Crazy' deception and lying as drama. *Anthropology & Medicine, 9*(2), 135–151.

Van Swol, L. M., & Braun, M. T. (2014a). Channel choice, justification of deception, and detection. *Journal of Communication, 64,* 1139–1159.

Van Swol, L. M., & Braun, M. T. (2014b). Communicating deception: Differences in language use, justifications, and questions for lies, omissions and truths. *Group Decision and Negotiation, 23*(6), 1343–1367.

Wood, L. A., & Kroeger, R. O. (2000). *Doing discourse analysis: Methods for studying action in talk and text.* Thousand Oaks, CA: Sage.

Yang, Y., Guyll, M., & Madon, S. (2017). The interrogation decision-making model: A general theoretical framework for confessions. *Law and Human Behavior, 41*(1), 80–92.

Contexts of Deceptive Communication: Interpersonal Relationships

Unchallenged Deceptions in Social and Professional Relationships

David Shulman

Introduction[1]

Self-deception occurs when individuals convince themselves that a false belief that they have is actually true (von Hippel & Trivers, 2011). Denial happens when, despite incontrovertible evidence, people reject a true state of affairs. A related behavior is encapsulated by the Russian word *Vranyo*, which describes social interactions where a speaker lies, an audience realizes that they are being lied to, and both parties pretend that nothing deceptive is happening. Here, the audience does not convince themselves that a false claim is true; instead, they pretend to believe the untrue claim. As in denial and self-deception, truth ends up taking a backseat to falsehood.

Unchallenged deceptions occur across a range of professional and social relationships. A manager will lie to a subordinate who pretends to believe the lie. A deceptive compliment about someone's physical appearance will go unchallenged by a willingly fooled flattered party. Family members and friends will not openly dispute a relative's deceptive claim to not drink or have financial problems. From 1993 to 2011, the US Department of Defense implemented the "Don't ask, don't tell" (DADT) rule regarding military service by gay and lesbian service members, making an official policy out of not challenging deception by omission. Interdisciplinary researchers have identified myriad examples of deceptions going unchallenged, such as Madoff's Ponzi scheme (Gibson, 2014), open secrets (Costas & Grey, 2016; Ledeneva, 2011), and in parenting children (see Goleman, 1985; Heyman,

D. Shulman (✉)
Lafayette College, Easton, PA, USA
e-mail: shulmand@lafayette.edu

Luu, & Lee, 2009; Laing, 1998). Table 27.1 provides some miscellaneous examples of unchallenged deceptions to orient the reader to this chapter's subject matter. More examples follow in the body of the chapter.

In these unchallenged deceptions, audiences believe that they can detect lies accurately. They register no uncertainty about whether they are being lied to. However, they are wary of the consequences of accusing a speaker of deception, so they pretend to accept false claims. When deceptions are identifiable, why do audiences choose not to accuse speakers of lying? David Gibson (2014) has stated in uncovering lying and secrets that it is "one thing to wonder, another to suspect and still another to accuse" (p. 293). This chapter explores why people hesitate "to accuse" by presenting five reasons

Table 27.1 Miscellaneous examples of unchallenged deceptions

Examples from scholarly research on unchallenged deceptions	Informal examples of unchallenged deceptions
Bribery to overlook Bernard Madoff's deception (Gibson, 2014)	A sister doesn't spoil her sister's false story to their parents about where she was last night in exchange for a future favor
Conformity to not disrupt the definition of the situation (Goffman, 1959, 1962, 1963) in face-to-face interactions, asylums, and stigma management.	Some older work colleagues dye their hair to appear younger. Their co-worker does not comment on the abundance of men and women over 60 with tar black hair and no hint of grey
Espionage (Goffman, 1969)	Spies let the deceptive cover stories of other spies pass without confrontation
Political subordination under powerful social institutions (Scott, 1990)	Accepting deceptive "official" version of reality. The old Soviet joke: "They pretend to pay us, we pretend to work" US Department of Defense "Don't ask don't tell" policy
To sustain workplace image, routine, and hierarchy (Dalton, 1959; Hughes, 1984; Jackall, 1988)	A manager makes a deceptive claim in a meeting. Workers who know better at the meeting pretend that the false claim is true A manager takes credit for work that a subordinate did and the subordinate does not challenge that allocation of credit
Family mandates pretense (Goleman, 1985)	Parents accept a child's lies about drug use
White lies (Ekman, 1985; Goleman, 1985)	The person complimented by a white lie knows that the compliment is untrue and accepts it anyway
Face-saving (Goffman, 1959)	Someone has shiny white caps on her teeth from undergoing a dental procedure. A co-worker knows that this visible degree of change could only come from having a procedure and asks the person about the experience. The person denies having had a dental procedure. The co-worker accepts that such dramatic change could come from "better brushing"

why people leave deceptions unchallenged: conforming by helping people sustain normal appearances, power inequalities, relationship maintenance, burden of proof issues, and pursuing individual ambitions.

Research on deceptive communication often focuses on problems associated with detecting whether a person is lying, appraising typologies of deception, and examining issues such as accuracy differences in detecting deception, mechanisms for improving deception detection, and physiological correlates of lying (for an overview of this literature see Granhag, Vrij & Verschuere, 2015). However, scholars have not investigated, to the same extent, audience reactions to detected deceptions, in terms of when they choose a strategic gullibility over confrontation. An impact of this research gap may be that people appear worse at detecting deception than they really are, with their reluctance to confront lies being mistaken for credulity.

von Hippel and Trivers (2011) argue that the research literature on detection of deception underestimates people's ability to detect lies. They criticize an overreliance on research in which: "(a) the deception is of little or no consequence, (b) the deceived has no opportunity to cross-examine the deceiver, (c) deceiver and deceived are strangers to each other, and (d) there are no repeated interactions between deceiver and deceived" (von Hippel & Trivers, 2011, pp. 3–4). The flip side of the four characteristics that von Hippel and Trivers identify—consequential deceptions, the capacity to probe deceptions, familiarity, and having repeated interactions—all, connects to sociological dimensions in naturalistic settings that help explain why people may avoid challenging deceptions.

von Hippel and Trivers (2011) also report that audiences do not reveal doubting deceivers even when they do. They speculate that deceivers may think that they are more successful in fooling others, and that people will pretend to believe deceivers to secure an information advantage or to keep the peace. Vrij (2011) disputes von Hippel and Trivers' (2011) methodological criticisms and argues in support of experimental findings that people are actually not that good at detecting attempts to deceive them. von Hippel and Trivers (2011) propose that scholars of deception and self-deception should investigate how "successful people are at detecting important deceptions occurring in naturalistic settings that allow those who are being deceived to gather further information as they see fit" (p. 4). Here, they suggest attending to an audience's strategic reaction to deception. This chapter acknowledges that the debate over people's genuine accuracy and skill in detecting lies is not yet settled. However, even if audiences are in error, the focus here is on cases where audiences feel positive that a speaker is lying.

Exploring unchallenged deceptions requires attending to the audience's experiences of reacting to lies. Many analytic concepts applied to lie-tellers could be amended to understand how audiences lie in not recognizing deceptions that they have detected. Concepts in the scholarship of lying can be revisited to fit this emphasis. Paul Ekman (1985) coined several conceptual

terms in the scholarship of lying, such as *detection apprehension* (fear of being caught in a lie), *duping delight* (joy felt from deceiving others), and *deception guilt* (guilt felt from deceiving others). Detection apprehension, for example, also emerges in the concerns and stresses that an audience can feel in even having to determine if deception is present. Deception guilt can exist as guilty knowledge in an audience, just as in a deceiver. Suppose a person knows of deceptions by others that end up victimizing third parties who the audience failed to alert? People can harangue those who did not challenge a deception that they knew of earlier, with a distressed, "why didn't you tell me what was going on?" Duping delight from deceiving others also exists in congratulating oneself on not being fooled and in secret pleasure in having an information advantage from discerning a lie.

Applicable Categories of Lies and Detection Costs

Audiences can encounter a variety of lies, including barefaced lies, distortions, omissions, half-truths, high-stakes lies, paltering (using truthful statements in an effort to mislead others), and white lies (for descriptions see Ekman, 1997; Goffman, 1959; Peterson, 1996; Rogers et al., 2017). Lies have antisocial and pro-social variants, both of which can go unchallenged. Ekman (1997) lists nine reasons why people lie: to avoid punishment; to obtain a reward not otherwise readily obtainable; to protect another; to protect oneself from the threat of physical harm; to win social admiration; to avoid an awkward social situation; to avoid embarrassment; to maintain privacy; and to exercise power over others by controlling the information that the audience has.

Ekman's nine motives for why people lie can reduce down to seeking to avoid costs associated with being honest and evade punishments in the form of reputational, legal, financial, and psychological distress. The act of lying can also pass on some of those costs to taken-in audiences. For example, some liars resentfully deny having gambling or drug problems, all while they steal from friends and family. Audiences can risk serious costs for not detecting hidden exploitative deceptions, such as adultery, financial frauds, cons, or a liar's incompetence in conducting meaningful duties. Audiences can also encounter another type of detection cost in probing deceptions. They may have to purchase equipment, like polygraphs or technical means of surveillance or MRIs, attend advanced training workshops, or employ third-party investigators, like hiring private detectives.

Perspectives on detection costs usually emphasize the costs of not detecting lies. This chapter uses the phrase detection costs in a different sense, as a catchall term to describe the unwelcome costs that an audience fears will emerge from actually detecting and accusing another party of lying. Deceivers can retaliate aggressively when they discover that their claims are being disputed. A quick example is evident in how whistleblowers suffer when they actually challenge deceptions and denials by superiors and peers at work,

instead of choosing to become silent colluders and pretending that nothing untoward exists.

Detection costs are a product of social context and not only the result of pure individual pathology or amorality. They encompass social and relational considerations that influence whether to ignore or confront perceived deceptions. Just as routinized pressures to lie exist in social groups, including in some families, close relationships, and workplaces (Goleman, 1985; Hunt & Manning, 1991; Jackall, 1988; Shulman, 2007), social norms also produce predictable pressures to not challenge some deceivers and their lies.[2] In this way, *Vranyo* is a structured social interaction.

Sociological Theory and Unchallenged Deceptions

To sociologists, analyzing social context is integral to understanding deceptive behavior. Bringing in the social context means attending to how large social forces such as stratification, class, race, gender, law, and economy are manifest in everyday situations in ways that influence individual actions. People leave deceptions unchallenged because socialization encourages conformity, power inequalities impose silence on audiences, burden of proof issues produce heavy reporting obstacles, people hope to maintain relationships, and individual ambitions can motivate pretending that no lying exists. These reasons connect to various social practices, institutions, and group norms.

For example, issues in arbitrating truth and falsehood are dictated not only by individuals but also by bureaucratic rules for negotiating conflicts and informal and formal sanctions. Burden of proof issues mean that if there is no convincing, standardized, or pain-free way for a counterclaim against a lie to move forward, people are left in a frustrating and stressful, "your word against mine," type quandary. The goal of maintaining individual relationships by overlooking deceptions is a bonding mechanism. Many social groups have a sacrosanct value that members do not divulge secrets. Social groups punish non-conformists who violate the status quo. Even though the value of truth-telling and shaming liars is preached as an ideal, people who tell are often shamed as "rats" or disbelieved. Powerful people punish those who cross them. Pursuing individual ambition and profit also comes from being part of a larger social system of rewards that flow depending on the structured incentives that institutional actors set.

How have sociologists examined the social contexts associated with these detection costs that can lead people to not challenge deception? Sociologists collectively have not written on deception to the extent that anthropologists, behavioral economists, communication scholars, criminologists, marketing and organizational behavior scholars, philosophers, and political scientists have. As a sociologist, this lag is distressing. Sociology, with its focus on theorizing and researching the behavioral influences of social contexts, can add

thoughtful perspectives to complement the insights other disciplines offer into deceptive behavior.

As background, sociological work on deception and lying frequently examines deceptions in naturalistic settings, ranging from criminological foci such as policing and financial crimes to workplace ethnographies and studies of distinct professions and groups and in close relationships (Anteby, 2008; Blum, 1994; Dalton, 1959; Gibson, 2014; Glaser & Strauss, 1964; Goffman, 1959, 1961, 1962, 1963, 1969, 1971, 1974; Hughes, 1984; Hunt & Manning, 1991; Jackall, 1988; Jacobs, 1992; Katz, 1979; Marx, 1988; Ruane, Cerulo, & Gerson, 1994; Scott, 2012; Shulman, 2007). This fieldwork focus supplies naturalistic venues of deception to complement the experimental settings that most other disciplines use. Sociologists also have contributed well-known interdisciplinary meta-overviews on deception (Barnes, 1994; Harrington, 2009). Finally, two preeminent sociological theorists, Georg Simmel (1950) and Erving Goffman (1959, 1963, 1969), are widely cited across diverse disciplines for their theorizing on deception.

A quick background in sociological assumptions is useful before diving into specific theories. Sociologists research how social forces such as class, gender, race, social control, power, and stratification influence people's actions and group outcomes, including deceptive behavior. Micro-sociological researchers, who work at the individual and small group levels of analysis, focus on how larger social institutions manifest those influences in face-to-face interactions. People are born into a world with preexisting social meanings. Individuals grow to learn definitions of how to act in social situations. Hewitt and Shulman (2011) describe the "definition of the situation" as "an organization of perception in which people assemble objects, meanings, and others, and act toward them in a coherent, organized way" (p. 49). Those definitions constrain people's behaviors into a conformity that guides their social actions. Social institutions such as class, family, and religion exert pressure in defining how people are to act within their social interactions.

People learn to enact stable sets of required performances that reflect their integrating social norms and values into their conduct. These stable performances connect to named statuses in society, with expectations of how to behave in a given status referenced as fulfilling a role. For example, a job title is a status. The appropriate actions someone in that status enacts are labeled a "role." Roles and statuses connect individuals to larger groups, organizations, and institutions, as roles and statuses are also organizational and institutionally dictated phenomena. Collective life, as Durkheim (1984) theorized, hits a key functionality when culture and conduct are in alignment and maintain a functional equilibrium. So, social groups and organizations tend to inculcate adherence to common interpretations for how to act in different situations. That inculcation should lead to predictable social behavior and produce social order. Social order will exist when stable, predictable, repeated behaviors, interactions, and interrelationships exist among individuals and within social

institutions. Sociologists study how that social order arises and is perpetuated and who wins and loses in its emergent arrangements.

CONFORMITY

How does the general sociological perspective presented above connect to lying? From a sociological vantage point, people are not always being "deceptive" in circumstances where they do not act authentically or truthfully like "themselves." While people ideally should be "themselves" in situations, individual autonomy is often curtailed by social expectations. For example, an individual boy or girl might want to act in a particular way, but forswear that autonomy to conform to gender expectations such as "boys don't cry" or being "ladylike." How people could act differs from how they are supposed to act when they inhabit specific roles and social statuses. While the individual is a decision-making unit in determining whether to challenge a lie or not, sociologists pay careful attention to what role that person is in and, in turn, what effects role requirements, social class, and other extra-individual factors have in whether a person like her or him should challenge a particular deception.

As an example, according to the historical US Department of Defense's DADT policy, being a gay man or woman was considered incompatible with soldiering. Hence, those in the role of soldiers had to deceive others about their sexuality, and in turn, those soldiers depended on others to not challenge those deceptions. A gay soldier could challenge the deceptive perception that military service and homosexuality are incompatible, and many did. But others did not, for fear of ostracism and putting their jobs at risk. For them, conversations about boyfriends and girlfriends who did not exist may have peppered conversations, and those wise to the secret, or who did not want to investigate the secret, asked no challenging questions.

Social life is replete with occasions where people learn not to act in any ways that could disrupt a definition of the situation. A person has obligations to join and enter the social scene before him and to help produce a scene collectively (the "happy" family or "team" at work) (see Goleman, 1985) by not acting to ruin that show. Abstaining from acting as an individual might prefer can lead to repressing some legitimate anger. Yet the expectation to conform or face the consequences is a powerful means of social control.

Erving Goffman is a preeminent theorist of dramaturgy, impression management, and face-to-face social interaction. Impression management refers to actions that individual actors take to present information about themselves to influence the perceptions that audiences will form about them. Goffman (1959) notes that when people convey a particular impression, the audience is to understand that a specific definition of the situation exists and that "a given state of affairs obtains" (p. 6). Here begins the pressure of socialization into conformity. People consciously work to manage social interactions so as to meet normative expectations. Their actions extend beyond honing

their own individual performances, to acting as part of a social group that sustains collective social images. People learn to overlook discrepancies between appearance and reality on the part of others in order to help sustain a "working consensus" in social life (Goffman, 1959, p. 6). They perform, both as individuals and in teams, to support those definitions of the situation. Goffman (1959) explains:

> Participants contribute to a single overall definition of the situation which involves not so much a real agreement as to what exists but rather a real agreement as to whose claims concerning what issues will be temporarily honoured. Real agreement will also exist concerning the desirability of avoiding an open conflict of definitions of the situation. Let us refer to this level of agreement as a 'working consensus.' (p. 6)

Socialization teaches people to support the working consensus against definitional disruptions. Goffman (1959) argues that people act to defend their own self-presentations against disruptions by engaging in defensive practices, one of which is lying. When individuals try to preserve a definition of the situation that another party offers, Goffman (1959) describes them as engaging in "protective practices." Not challenging deceptions fits within the protective practices concept in maintaining a working consensus. People are to conform to the normal appearances associated with a definition of the situation. They are to help sustain those normal appearances, even when those appearances are deceptive. Examples of this pressure are evident in many subsequent examples of unchallenged deceptions.

Power Inequalities

The respective social positions of audience and speaker impact an audience's decision to confront a deceiver. People can have a larger physical size or a weapon to force agreement with deceptive claims. But power here also involves set roles and social circumstances that come with resources that can pressure an audience to meet particular expectations. Those forces are not properties inherent to the individuals themselves, as physical size and fists are, but are drawn from resources associated with a role and status that are held to be legitimate in social situations. For example, supervisors have the capacity to damage a subordinate's employment status, typically without reference to whichever specific individuals are in that power relation. Authority and hierarchy are social forces that can be drawn upon in encounters between superordinate and subordinates.

A relevant example comes from a friend who complained about a high-level manager who told an assembled group of co-workers, "Titles don't matter. We are all one team." My friend resented this remark as being knowingly deceptive and manipulative, as titles dictate pay, and there is no "one team" philosophy in place when allocating salary. The egalitarianism disappears

when deciding on compensation. The manager's deception that "titles don't matter" illustrates a powerful person stating a definition of the situation that was meant to motivate solidarity and hard work and also to be uncontested. The high-level manager has a title and associated power that exempted her claim from a certain level of "reality check" skepticism.

Similarly, students tell stories from their youth sports days of coaches claiming that everyone was important in the team's victory, implying that those on the bench are equal to those who start the games. Yet when playing time is mostly allocated to starting players, and only rarely to others, all team members could not possibly be equally influential partners in victory. The students concluded that the deceptive claim of everyone being equally important was intended to stifle concerns over playing time, including shaming individuals as "bad team players" if they complained. In both situations, there is power inequality and a superordinate making problematic claims about what the true situation is to subordinates. Respondents are supposed to agree, or at least not openly contradict that claim.

Those with more powerful roles in social situations have a higher expectation that their claims about reality will be accepted and any skepticism about them hidden. James Scott (1990) coined the term *public transcripts* to describe definitions of the situation that dominated parties accede to in public to suit powerful individuals who oppress them. Public transcripts comprise the "official version" of situations. James Scott (1990) uses the term *hidden transcripts* to refer to the private disagreements and consequent interpretations of those prevailing definitions that find expression offstage, as they cannot be directed publicly against the powerful. Scott is interested in how hidden transcripts form a path of resistance that weaker parties have against more powerful ones. Hidden transcripts offer an information-advantage and pride-salvaging process that comes from not challenging deceptions openly and hiding an actual resistant consciousness. They allow people a chance to recover somewhat from the subordination involved in not challenging a powerful person's misrepresentations to their face, and to remind themselves and others, that "in reality," they do see the truth. It is likely a rare employee who has not encountered a situation yet in which they felt pressure to honor a superior's rationalization, despite their own private disagreement.

Working consensus also produces an expectation that one will lie on behalf of others as well as not notice their lies. Such collusion requires more commitment than just not challenging some lies in order to maintain an advantageous definition of the situation. Robert Jackall (1988) summarizes these kinds of expectations for aspiring managers in his ethnography *Moral Mazes*:

(1) You never go around your boss. (2) You tell your boss what he wants to hear, even when your boss claims that he wants dissenting views. (3) If your boss wants something dropped, you drop it. (4) You are sensitive to your boss's wishes so that you anticipate what he wants; you don't force him, in other words, to act as boss. (5) Your job is not to report something that your

boss does not want reported, but rather to cover it up. You do what your job requires, and you keep your mouth shut. (pp. 109–110)

While identifying a liar might be virtuous, there is tremendous social pressure against being a "tattletale." Being a "rat" violates the everyday protective practices that individuals enact to avoid acknowledging discrepancies in the images of others present. Far from just being an individual-level phenomenon of protecting others, there are explicit group-level rewards for hiding the deceptions of others within that group. For example, one benefit of being ensconced in an organization or a group is that members join together to face external threats (Coser, 1956). If group members think of other people in the group as being untrustworthy, that action threatens the group's functional unity. The full power of the group's displeasure may then fall upon the discrepant member.

People are socialized to learn to pay civil inattention to disruptive information (Goffman, 1959). Paul Ekman (2009) speculated that some inaccuracy in lie detection is attributable to lag effects from parents teaching children not to detect lies. In this vein, Susie Scott notes that people pretend not to know that they are to overlook information that would disrupt definitions: "Feigning ignorance of their mutual understanding that things are not as they appear: the 'elephant in the room' is tactfully overlooked" (Scott, 2012, p. 265). Scott (2012) argues that people "are… given a distorted version of the truth but they raise no objections because they recognize its social value" (p. 266).

Scott identifies examples of these unchallenged deceptions in analyzing gendered divisions of labor within relationships. Men and women will offer up stories that gender preferences explain unequal work in the home (women want to cook and men don't care about messes), which rationalizes an inequality that could otherwise spoil the relationship. Scott (2012) uses Goffman's idea of transreceivership to understand people's ability to be aware of both the role of the deceived and deceiver, so that they see from the perspective of each. In her words, we "perceive our teammates as individualist actors who are as shrewd, calculating and instrumental as we are ourselves" (Scott, 2012, p. 272).

Individuals enter social situations preprogrammed with instructions about the kinds of working consensus to maintain there. Deceptive claims within some of those spheres, if enacted to support the definition in power, are to remain unchallenged. Robert Jackall (1988) quoted a respondent telling him: "What is right in a corporation is not what is right in one's home or church. What is right in the corporation is what the guy above you wants from you" (p. 164). Truthfulness and falsehood become less absolute and more morally ambiguous, as real-life contexts introduce vested interests in how "facts" are understood. Jackall (1988) comments on this transition for managers: "'truth' is socially defined, not absolute, and… therefore compromise, about

anything and everything, is not moral defeat… but simply an inevitable fact of organizational life" (p. 111).

Ledeneva (2011) defines an open secret as "unarticulated knowledge that everybody who is party of a transaction knows about but that no one discusses in a direct way" (p. 725). Sometimes people are tested to show that they know the importance of not noticing or appearing to know as a form of expressing a dominant power relationship. Ledeneva offers this anecdote of an audience member who did not take this lesson into account with his political superiors, and was scolded to do so:

> A politburo member is giving a speech about industrialization and twenty-story skyscrapers recently built on Karl Marx Street in Kharkov. Suddenly one of the listeners interrupts him:
>
> "Comrade Kalinin, I am from Kharkov. I walk down that street every day, but I have not seen any skyscrapers!"
>
> "Comrade," replies Kalinin, "instead of loitering on the streets you should read newspapers and find out what's going on in your city."[3]

Numerous commentators on workplaces note that unwritten rules and informal sets of practices govern how people really get their work done. Ledeneva (2011) states, "unwritten rules define the ways of circumventing constraints, both formal and informal, of manipulating their enforcement to one's own advantage, and of avoiding penalties" (p. 722). Unchallenged deceptions occur in not acknowledging the actual workings of the system out loud to preserve a power to make use of hidden, informal work practices (Dalton, 1959; Goffman, 1963). For example, say a due date is distributed for departments to turn in a burdensome report to administrators. Experienced workers know from past years that the stated deadline is a "soft deadline," and that they can secretly take more time to turn in the work. Yet the experienced workers do not spoil the show of the "deadline" by publicly correcting the date, which adheres to the unwritten rules. Secret flexibility is to be unpublicized lest it be lost.

Sociologists have examined lying as a "micro-political maneuver of knowledge and power connected to knowledge" (Rodriguez & Ryave, 1990, p. 196). Such maneuvering is present in the face-work (Goffman, 1967) and protective practices (Goffman, 1959) that preserve another's image in service of etiquette and politeness. While lying can be spontaneous in individual interaction, having to tell particular lies repeatedly may form predictable routines in workplace settings (Shulman, 2007). Rodriguez and Ryave (1990) argue that liars experience an organizationally motivated basis for lying in preferring social acceptance to being rejected. Routine lies reflect institutional practice. Not challenging them is also a repeated requirement. Co-workers, for example, may not challenge a deceptive claim by another

worker to preserve face for the profession (Hughes, 1984). These deceptions can be motivated for reasons other than to save one's own or a colleague's skin. They can attempt to preserve a veneer of expertise, power, and authority that is externally accorded to the profession, and onlookers may support that mission enough to not correct the record. Unchallenged deceptions can also occur to save other people's skins, as when some police officers at crime scenes do not challenge false claims by other arresting officers of the conditions during an arrest. These incidents usually connect to a developed "us versus them" mentality associated with the "thin blue line" (Klockars, 1985) and also comprise a form of relationship maintenance.

Relationship Maintenance

Daniel Goleman (1985) notes in his analysis of self-deception that "the tradeoff between a distorted awareness for a sense of security is, I believe, an organizing principle operating over many levels and realms of human life" (p. 21). He argues that people engage in self-deception by dimming their awareness to enhance their sense of security. People may elect not to challenge deceptions that, if confronted, could dissolve an important relationship. People deceive in intimate relationships as a conflict avoidance technique (Peterson, 1996); they also do not challenge some of a partner's deceptions as a conflict avoidance technique. Cole (2001) found support for his hypotheses that among close relationships, being successfully misled is related to satisfaction with one's partner, and also that fear of abandonment is related to using deception.

In a different study, 325 respondents were asked about a time when they knew that someone in a close relationship lied to them in an impactful way (Grayson & Shulman, 2017). In a significant subset of cases, the respondents maintained a close relationship with persons who deceived them. They did not confront the deceiver, try to repair the trust they had in the relationship, or tell other people about their partner's lies. Even knowing that they were being lied to, they choose to maintain the relationship status quo by not challenging the deceiver. When asked why, some noted that the relationships that they were in also intertwined with relationships with others. A relation with a bandmate who lied didn't just imperil the relation with the person but also with the band. Confronting problems with someone in a family or in a group of friends risked the respondent's connections to other people as well. An audience faced not only the prospect of costs in the immediate relationship with the deceptive partner; they could face higher costs when that relation overlapped with other relationships that would be affected if the partner's deceptions were challenged. In this vein, Rivera (2012) argues that the cultural fit between employees and potential hires is vital. One means of fitting in is to understand what working consensus a given group prefers to maintain in different situations, including knowing that working consensus can be sustained by not challenging deceptions. Not challenging deceptions

can demonstrate one's fit with the cultural norms of a group and can show partners to be kindred spirits.

People also avoid challenging deceptions when those deceptions can tell them what they want to hear about themselves. While scholars write about the pro-social nature of white lies and the value of not hurting people's feelings, there is also an unchallenged deception that can lurk in the recipients accepting the white lie as truthful when they know otherwise. The debate over the morality of white lies is usually about the teller, rather than about the receiver, who willingly accepts false information and has people collude in sustaining that illusion.

Glaser and Strauss (1964) identified different kinds of awareness that people display in social interaction. They identify a "pretense awareness context" in which interactants are mutually fully aware of true and false claims but "pretend not to be" (Glaser & Strauss, 1964, p. 670). Glaser and Strauss conducted their ethnographic work in a hospital researching how medical staff, patients, and their families interacted around mortality. The pretense awareness context emerged "when patient and staff both know that the patient is dying but pretend otherwise—when both agree to act as if he were going to live—then a context of mutual pretense exists" (Glaser & Strauss, 1964, p. 64). In the pretense awareness context, both patient and staff are aware of the actual facts but they pretended otherwise to smooth over the difficulties of the situation. This might be considered a routinized white lie. Compassion is long-held as a motive for pro-social deception and may also be a contextual force in not challenging some deceptions. A mutual pretense awareness context also requires policing interactions for any slips that could expose the pretense (Glaser & Strauss, 1964).

Not challenging deceptions also can lessen how much work people have to do. Blum (1994) found lying by caregivers to be endemic in dementia care settings to avoid agitating patients. Many deceptions go unchallenged in front of those who know the truth in order to expedite work and not create conflict in a working relationship.

BURDEN OF PROOF ISSUES

An audience seeking to challenge a deceptive statement will identify and apply some evidentiary standard when they confront the liar. The ensuing confrontation may be limited to the audience and the teller. However, a heated conflict always runs the risk of being exposed to other people. A person seeking to challenge a lie can fear that any larger audience to the dispute can turn against them. When third parties enter the picture, their alliances and priorities may pose a problem for the person who confronts deception. The evidentiary case that the challenger has may not meet a burden of proof that satisfies others. A bureaucratic process of hearing out a challenge can be so onerous in itself that potential challengers would rather live with an unsanctioned lie than be punished by a difficult procedure. The employee handbook

connected to challenging the relevant type of deception may involve long drawn out procedures, forbid anonymity, require face-to-face confrontations and cross-examinations, and not safeguard challengers enough from retaliation. This problem is endemic in pursuing some crimes. Numerous misdeeds and lies also go unreported for fear that allegations of them will not be believed, and create new problems for the challenger. The response of "are you calling me a liar" almost always doubles as a pretext to go after a challenger in the severest way possible. The process of challenge and establishing that the other person lied can deter confronting a liar.

Individual Ambition

Individual ambition can lead to tolerating deception as a necessary evil. The ends justify the means is a familiar justification of lies. Bribery is an example of not challenging deception that constitutes a form of incentivized, insincere gullibility. In the Madoff case, his deceptions went unchallenged by some who profited by leaving them uninvestigated (Gibson, 2014). A more down-to-earth example is the sibling who does not rat out another sibling about where he or she really "studied last night" for some payoff. Having knowledge of someone having lied is also a resource that people retain for later use in more advantageous circumstances. Like the spy who knows a person's secret, a decision to go public with the detected lie can be a matter of waiting for the best time for exposure.

The workplace, as Jackall's (1988) work clarifies, also rewards those who do not challenge lies. They can show themselves to "be one of us" and gain further favor down the road. Not challenging a flattering white lie is also in a person's self-interest, in hearing self-esteem-boosting information. Not dissolving a useful relationship, even in the face of deception, may be considered a self-interested decision. Sometimes someone's economic or psychological security is connected to that relation and would be lost. Greed, upward mobility, personal security, and relief from having to deal with the stress of mounting a challenge are all forms of self-interest that help explain not challenging deceptions.

Directions for Additional Research

The subject of unchallenged deceptions offers much unexplored territory for additional research. Some key questions to explore in future work include:

- Scholars of deception have identified various kinds of lies. Do audiences tend to confront some kinds of lies more than they challenge others?
- How do the demographic makeups of a liar and a recipient of a lie connect to a higher "challenge rate," namely a tendency to challenge more or less depending on who each party is (gender, race, age, and class)?

- Is actual effectiveness in detecting deception artificially lower because of people's reluctance to confront deception? Does a high and accurate rate of stealth detection exist?
- Levine (2014) proposes a truth-default theory, arguing that deception is relatively infrequent and that a truth-default approach has adaptive qualities that are useful socially. The truth-default theory has more difficulty in explaining unchallenged lies that help small groups and workplaces function more effectively. More productivity may come from embracing collective fictions than truths. The intersection between unchallenged deceptions and truth-default theory bear more examination.
- Sociologists interested in dramaturgy should explore the actions involved in unchallenged deception more. A liar is interested in noting how the audience responds to the lie. Does a person believe what they hear? Will the liar get away with the deception? In unchallenged deceptions, there is impression management work to carry off pretending to believe well enough to avoid an unwanted response by the liar. So how does *Vranyo* work as a performance by an audience?
- When do deceivers know that audiences have accurately detected their deceptions versus when does the audience's pretense fool them?

Notes

1. "Deception" and "lies" are used interchangeably here. While many papers make definitional distinctions between types of deception and lying, they are treated here as synonyms for when people willfully communicate untrue information without prior notification of intent to deceive (Ekman, 1985).
2. An interesting research question to pursue in this regard is when the "who" matters more for why the lie is not challenged than the lie's content does. A powerful person can command involuntary gullibility regardless of a lie's subject. In other circumstances, the severity and stakes of the lie may have more impact on challenging it.
3. Ledeneva cites the source of this anecdote as from Seth Graham. (2004). "Varieties of reflexivity in the Russo-Soviet anekdot", p. 176. In L. Milne (Ed.), *Reflective laughter: Aspects of humour in Russian culture* (pp. 167–180). London: Anthem.

References

Anteby, M. (2008). *Moral gray zones: Side production, identity and regulation in an aeronautic plant.* Princeton, NJ: Princeton University Press.

Barnes, J. A. (1994). *A pack of lies: Towards a sociology of lies.* Cambridge: Cambridge University Press.

Blum, N. S. (1994). Deceptive practices in managing a family member with Alzheimer's disease. *Symbolic Interaction, 17*(1), 21–36.

Cole, T. (2001). Lying to the one you love: The use of deception in romantic relationships. *Journal of Personal and Social Relationships, 18,* 107–129.

Coser, L. (1956). *The functions of social conflict.* Glencoe, IL: The Free Press.

Costas, J., & Grey, C. (2016). *Secrecy at work: The secret architecture of organizational life.* Stanford, CA: Stanford University Press.

Dalton, M. (1959). *Men who manage.* New York: Wiley.

Durkheim, E. (1984). *The division of labor in society.* New York, NY: The Free Press.

Ekman, P. (1985). *Telling lies.* New York: Norton.

Ekman, P. (1997). Deception, lying and demeanor. In D. F. Halpern & A. E. Voiskounsky (Eds.), *States of mind: American and post-Soviet perspectives of contemporary issues in psychology* (pp. 93–105). New York, NY: Oxford University Press.

Ekman, P. (2009). Lie catching and micro expressions. In C. Martin (Ed.), *The philosophy of deception.* New York, NY: Oxford University Press.

Gibson, D. R. (2014). Enduring illusions: The social organization of secrecy and deception. *Sociological Theory, 32,* 283–306.

Glaser, B. G., & Strauss, A. L. (1964). Awareness contexts and social interaction. *American Sociological Review, 29,* 669–679.

Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday.

Goffman, E. (1961). *Encounters: Two studies in the sociology of interaction.* Indianapolis, IN: Bobbs-Merrill.

Goffman, E. (1962). *Asylums: Essays on the social situation of mental patients and other inmates.* Chicago: Aldine.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity.* Englewood Cliffs, NJ: Prentice-Hall.

Goffman, E. (1967). *Interaction ritual: Essays in face-to-face behavior.* Doubleday Anchor.

Goffman, E. (1969). *Strategic interaction.* Philadelphia: University of Pennsylvania Press.

Goffman, E. (1971). *Relations in public.* New York: Basic Books.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience.* Cambridge, MA: Harvard University Press.

Goleman, D. (1985). *Vital lies and simple truths.* New York: Simon and Schuster.

Granhag, P., Vrij, A., & Verschuere, B. (2015). *Detecting deception: Current challenges and cognitive approaches.* Chichester: Wiley Blackwell.

Grayson, K., & Shulman D. (2017). *Et Tu, Brute*? Vulnerabilities to opportunism in relational embeddedness. Unpublished manuscript.

Harrington, B. (Ed.). (2009). *Deception: From ancient empires to internet dating.* Stanford, CA: Stanford University Press.

Hewitt, J., & Shulman, D. (2011). *Self and society: A symbolic interactionist social psychology.* New York: Pearson.

Heyman, G. D., Luu, D. H., & Lee, K. (2009). Parenting by lying. *Journal of Moral Education, 38,* 353–369.

Hughes, E. (1984). *The sociological eye.* New Brunswick, NJ: Transaction Press.

Hunt, J., & Manning, P. (1991). The social context of police lying. *Symbolic Interaction, 14,* 45–56.

Jackall, R. (1988). *Moral mazes: The world of corporate managers.* New York: Cambridge.

Jacobs, B. A. (1992). Undercover drug-use evasion tactics: Excuses and neutralization. *Symbolic Interaction, 15,* 435–453.

Katz, J. (1979). Concerted ignorance: The social construction of cover-up. *Urban Life, 8,* 295–316.

Klockars, C. B. (1985). The dirty harry problem. In F. A. Elliston & M. Feldberg (Eds.), *Moral issues in police work* (pp. 55–71). Totowa, NJ: Rowman & Allanheld.

Laing, R. D. (1998). *Selected works RD laing: Self & other V2*, Taylor & Francis.

Ledeneva, A. (2011). Open secrets and knowing smiles. *East European Politics and Societies, 25,* 720–736.

Levine, T. R. (2014). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392.

Marx, G. T. (1988). *Undercover: Police surveillance in America.* Berkeley: University of California Press.

Peterson, C. (1996). Deception in intimate relationships. *International Journal of Psychology, 31,* 279–288.

Rivera, L. (2012). Hiring as cultural matching: The case of elite Professional service firms. *American Sociological Review, 77,* 999–1022.

Rodriguez, N., & Ryave, A. (1990). Telling lies in everyday life: Motivational and organizational consequences of sequential preferences. *Qualitative Sociology, 13,* 195–210.

Rogers, T., Zeckhauser, R. Francesca G., Norton, M. L., & M. E. Schweitzer. (2017). Artful paltering: The risks and rewards of using truthful statements to mislead others." *Journal of Personality and Social Psychology, 112,* 456–473.

Ruane, J. M., Cerulo, K. J., & Gerson, J. M. (1994). Professional deceit: Normal lying in an occupational setting. *Sociological Forum, 27,* 91–109.

Scott, J. (1990). *Domination and the arts of resistance: Hidden transcripts.* New Haven, CT: Yale University Press.

Scott, S. (2012). Intimate deception in everyday life. *Studies in Symbolic Interaction, 39,* 251–279.

Shulman, D. (2007). *From hire to liar: The role of deception in workplace culture.* Ithaca, NY: Cornell University Press.

Simmel, G. (1950). *The sociology of George Simmel.* New York: Free Press.

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34,* 1–56.

Vrij, A. (2011). Self-deception, lying and the ability to deceive. *Behavioral and Brain Sciences, 34,* 40–41.

Angry Hugs and Withheld Love: An Overview of Deceptive Affection

Sean M. Horan and Melanie Booth-Butterfield

> *It's discouraging to think how many people are shocked by honesty and how few by deceit.*
>
> —Noël Coward, Blithe Spirit

Noël Coward's observation may explain why relational partners are not always authentic in their expression of affection to one another. It may be more expedient, simpler, or even strategic to adjust one's communication of affection to messages others expect to receive. This process is termed *deceptive affection*. This chapter examines the foundations and motivations for deceptive affection, discusses associated research findings, describes theoretical frameworks, and considers ways in which to extend the research exploring this process.

Overview of Construct

Research identifying and describing deceptive affection was born out of comparing affection exchange theory's (AET; Floyd, 2006) claims and existing deception research. AET argues that "Affectionate feelings and affectionate

S. M. Horan (✉)
Fairfield University, Fairfield, CT, USA

M. Booth-Butterfield
West Virginia University, Morgantown, WV, USA
e-mail: mboothbu@wvu.edu

535

expressions are distinct experiences that often, but need not, covary" (Floyd, p. 163). Deception research using diary methods revealed that individuals regularly lied to their non-married partners (DePaulo & Kashy, 1998), and feelings were a frequent topic (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996). The combination of AET's arguments and DePaulo's findings regarding feelings-based deception suggested that it was necessary to explore instances where *feeling* and *expressing affection* specifically were divergent. Consequently, Horan and Booth-Butterfield (2011, 2013) began research examining deceptive affection.

Fundamental to the study of deceptive affection is distinguishing between *felt* and *communicated* affection. Affectionate feelings entail "warmth and fondness toward someone" (Andersen & Guerrero, 1998, p. 59), which can be conveyed via affectionate communication: "the process of expressing our care, appreciation, value, and love for others" (Floyd, 2006, p. xiii). When the *feeling* and *expression* of affection differ, deceptive affection occurs (Horan & Booth-Butterfield, 2011). Deceptive affection can occur one of two ways. First, it can occur through active deceptive affectionate messages (DAMs). That is, DAMs can occur both when expressing affection when no affection is actually felt, and/or by intensifying experienced affection (e.g., feeling jealous or angry yet expressing affection instead of those negative feelings). Horan and Booth-Butterfield drew upon display rules (Ekman & Friesen, 1975) to explain this process: individuals can simulate or intensify affectionate feelings via communication.

The second way deceptive affection occurs is when individuals withhold expressing affection that they truly experience (e.g., feeling strongly about someone in the beginning stages of a relationship, but holding back expressing the true intensity of those affectionate feelings). Using display rules, and based on AET's (Floyd, 2006) framework, Horan and Booth-Butterfield reasoned that individuals could deintensify, mask, or inhibit felt affection. Thus, deceptive affection is consistent with definitions from preeminent deception scholars (e.g., Buller & Burgoon, 1998; DePaulo et al., 1996; McCornack, 1992; McCornack, Morrison, Paik, Wisner, & Zhu, 2014), such as: "a message knowingly transmitted by a sender to foster a false belief or conclusion by the receiver" (Buller & Burgoon, 1998, p. 381). However, it is important for scholars to distinguish general deception, which may be in reference to any topic, from deceptive affection, which focuses on the positive emotional, affectional aspects of a relationship.

DECEPTIVE AFFECTION: HOW AND WHY

Initial deceptive affection studies aimed to understand whether people could recognize deceptive affection and explain why they undertook such deceptive strategies. To that end, authors designed deception diary studies, based on DePaulo (DePaulo & Kashy, 1998; DePaulo et al., 1996), to investigate both

DAMs (Horan & Booth-Butterfield, 2013) and withheld affection (Carton & Horan, 2014). Both studies used the same general model: participant diary instruction and pre-survey measures, maintenance of a weeklong diary, and outcome measures.

In the first study, participants in non-married romantic relationships, lasting at least three months, maintained a 7-day diary (Horan & Booth-Butterfield, 2013). In these diaries, participants described the feelings they experienced, inauthentic affection they expressed, and their motives for expressing DAMs. Nearly all diaries contained DAMs and these were expressed an average of 3.30 times a week. The feelings lied about pertained to the message source (e.g., self-oriented), their relational partners, or the situational context. Example feelings included "jealous, annoyed, frustrated, regret, hung-over, exhausted" (pp. 206–207). Instead of expressing these negative feelings, participants instead elected to communicate affection.

DAMs were communicated both verbally and nonverbally. Verbally, communicators expressed confirming and/or avoidant responses of affection. Confirming examples "included compliments about appearance and statements of enhanced pleasure or joy about spending time together" (Horan & Booth-Butterfield, 2013, p. 207). Avoidant examples generally entailed those times when individuals expressed affectionate messages as a means of partner or topic avoidance. Nonverbally, DAMs were expressed via proxemics, haptics, and kinesics—essentially, nonverbal modes of affection such as kisses and hugs. Motives for DAMs were face-saving, conflict management/avoidance, and emotion management. For more information about the feelings lied about, DAMs used, and motives, readers are encouraged to review Horan and Booth-Butterfield (2013).

The previously summarized diary research described one mode of deceptive affection. Therefore, Carton and Horan (2014) undertook a similar study, this time aiming to describe how and why people withhold affection. Romantic and sexual partners in this study withheld affection 5.67 times a week. The most frequently experienced feelings included *liking* and a *desire for affection*. When reporting other feelings, the authors explained: "Other participants experienced less positive emotions when struggling with feelings and their decisions to withhold affection" (p. 228). These included: *irritation/anger*, *anxiety*, and *general contentment*. Instead of expressing their true affectionate feelings, they primarily reported expressing *no new action*, *moderated affection*, *nonchalance*, and *disaffectionate* messages. Commonly reported motives for withholding authentic affectionate expressions included *concern for perception*, *inappropriate circumstances*, *negative emotions*, and *fear of outcom*e. For a complete discussion of withholding affection, see Carton and Horan (2014).

These initial explorations demonstrate that instances of deceptive affection, via DAMs and withheld affection, were quantifiable, recognizable, and could be explained by their communicators. These findings underscore the

importance of studying the deceptive affection process. Subsequent research aimed to understand the implications of expressing deceptive affection.

CONSIDERING THE RISKS OF DECEPTIVE AFFECTION

Following the description of deceptive affection offered by diary-based research, additional research explored potential risks, or drawbacks, of engaging in deceptive affection. Both affectionate communication and deception are frequently described as *risky* (Floyd, 1997a, 1997b, 2001, 2006; Floyd & Voloudakis, 1999; Levine, McCornack, & Avery, 1992; Morman & Floyd, 1998; Rogers, Zeckhauser, Gino, Norton, & Schweitzer, 2017). There are possible pitfalls of both, expressing authentic affection to partners, or attempting to deceive them. Therefore, a key question cutting across initial studies was whether deception and affection combined present drawbacks or risks. Potential negative outcomes could include detection, negative emotional reactions within the source of the DAM, and increased psychological and physiological stress (Horan & Booth-Butterfield, 2011).

To better understand risks associated with deceptive affection, researchers examined the question of whether communicating inauthentic affection to a romantic partner was stressful. Specifically, Horan and Booth-Butterfield (2011) examined physiological and emotional implications of recalled deceptive affection. Based on prior affectionate communication experiments and emotional writing (Floyd, 2006; Pennebaker & Beall, 1986), participants were randomly assigned to write about either (1) a romantic partner DAM, (2) authentic partner affection, or (3) plans with a friend (control group). Blood pressure and heart rate were measured before and after the 20-minute writing exercise, to discern whether DAMs were distressful to participants. Through a variety of analyses, the collective results suggest that participants were not physiologically activated when writing about DAMs. Horan and Booth-Butterfield (2011) also created a measure to explore differences in outcomes based on motive for deception, but those results could be reconsidered given that an inductive typology of actual motives was later identified (Horan & Booth-Butterfield, 2013).

The motives for deceptive affection identified, combined with the physiological findings, suggest that deceptive affection is not significantly risky for sources. It may be a normal part of relational interactions. Of course, four factors should temper this claim. First, most deceptive affection research has focused on romantic relationships. Less is known about the potential problems posed by these messages in casual sexual relationships, one-night stands, on-again/off-again relationships, and friends with benefits. Second, these studies have explored what appear to be normative examples of deceptive affection (e.g., Carton & Horan, 2014; Horan & Booth-Butterfield, 2013). Future research should explore those instances of deceptive affection that resulted in internal/external conflict and/or those considered to

be turning points (Horan & Booth-Butterfield, 2011). Third, instances of deceptive affection used for anti-social purposes may be more problematic for sources and receivers. Deceptive affection used to manipulate another person could be troubling on many levels (see Floyd, 2006). Fourth, it is unknown whether one mode of deceptive affection (DAMs vs. withholding affection) has more significant implications in relational communication and/or is indicative of problems.

Deceptive Affection and Relational Maintenance

If deceptive affection with partners is not considered particularly risky, then how else might it be viewed? One proposed answer is relational maintenance (Horan, 2013b; Horan & Booth-Butterfield, 2013). Horan and Booth-Butterfield initially argued that DAMs could be employed to maintain a relationship, keeping it on a positive trajectory. This argument is consistent with prior studies positioning affection (e.g., Guerrero & Bachman, 2006) and deception as relational maintenance (Guthrie & Kunkel, 2013). It also aligns with AET's (Floyd, 2006) proposition that affection enhances pair bonds. Hence, partners may deceive about their level of affection to increase partner closeness and commitment in ongoing relationships.

As Horan and Booth-Butterfield's (2013) findings were argued to be evidence of DAMs as maintenance, Horan (2013b) specifically explored the relationships among deceptive affection and relational maintenance. He found that participants expressed DAMs to their romantic partners 3.64 times a week, and withheld affection from partners 3.66 times a week. Frequency of DAMs was positively related to the frequency of withholding affection within the relationship, suggesting an overall lack of affection related authenticity in some relationships. Further, DAMs were negatively related to assurances and positivity, whereas withholding affection was unrelated to prosocial forms of maintenance. In contrast, examining negative relational maintenance (Dainton & Gross, 2008), both frequency of expressed DAMs and withholding affection were related to allowing control. Future studies would be wise to more completely disentangle the relationships among authentic and deceptive affection and relational maintenance communication. This may be especially important as Horan noted that certain relational maintenance scale items directly implicate affection.

Given that reported motives are not overly negative, that deceptive affection does not seem distressful, and that maintenance has been a proposed argument for deceptive affection, some might view deceptive affection as selfless or altruistic deception. Deceptive affection researchers have noted this idea, drawing on Kaplar and Gordon's (2004) study of altruistic deception. Those researchers had participants provide narratives of lies they told former romantic partners, and analyses revealed the presence of selfish motives. Viewing Kaplar and Gordon's findings along with their own, Carton

and Horan (2014) reasoned that "sources have conflicting and often inaccurate views about their own deception" (p. 237). Consequently, though motives aimed at enhancing partner feelings, face-saving, and mood enhancement have been noted for deceptive affection it is apparent that some selfish motives are present. That is, participants may express DAMs to save their partner's face, but inherent in that would be a selfish motive as well, because a person would feel negatively for hurting his or her partner's feelings and conflict could result. Kaplar and Gordon were critical of deception diary designs, and future studies are encouraged to replicate their methods when studying deceptive affection. As they summarized: "Our study highlights that data obtained using different methods may differ in some respects. Researchers using questionnaire methods may obtain different results than had they used content analysis (and vice versa)" (p. 506).

Deceptive Affection and Additional Relationship Explorations

The body of deceptive affection research is young, yet researchers have aimed to understand factors that help explain this process. For example, Gillen and Horan (2013) were concerned with understanding how the frequency of DAMs related to relational qualities. Such an interest was governed by AET's argument that affection enhances relational qualities, and studies support such a claim (Floyd, 2006; Horan, 2012; Horan & Booth-Butterfield, 2010). Participants in their study reported expressing DAMs 3.49 times a week to romantic partners. Frequency of DAMs was unrelated to commitment and satisfaction, yet the frequency of general partner deception was negatively related to commitment and satisfaction. Though previous studies had examined motives for deceptive affection, Gillen and Horan extended this idea by exploring beliefs about deception (Scholl & O'Hair, 2005). Curiously, frequency of DAMs was unrelated to deceptive beliefs, yet general partner deception was negatively related to these beliefs. Gillen and Horan concluded: "It appears that communicators may view DAMs and deception differently" (p. 356). Future studies are encouraged to understand how people might view deceptive affection, general deception, and altruistic deception (if such a thing exists) differently.

   The majority of the studies reviewed have been concerned with describing deceptive affection, exploring risks, and understanding correlates that might help explain the frequency of communication. Departing from this perspective, Trask, Hortsman, and Hesse (2016) examined an aspect of deceptive affection across a variety of relationships. That is, they created a scale to measure deceptive affection and examined differences in the frequency of communication based on relationship type. Their findings revealed some differences in the frequency of deceptive communication and relational qualities among cross-sex friends, friends with benefits, and romantic relationships.

It is important to note, though, that there are concerns as to whether the scale items utilized represent the full range of deceptive affection. Upon initial reading of the scale's face, items might only capture intensified DAMs and deintensified (withheld) affection. Subsequent studies of deceptive affection should incorporate items that capture the full range of deceptive affection.

In addition to the previous deceptive affection studies, researchers have focused on one specific form of deceptive affection via manipulative affection. Floyd, Ebert, Davis, and Haynes (2005, as reported in Floyd, 2006) described manipulative affection as instances when individuals "expressed affection that they *did not actually feel…*for a manipulative purpose" (p. 137). Accordingly, manipulative affection is a specific type of DAM wherein inauthentic affection is expressed for a strategic objective. Consistent with general deceptive affection research, manipulative affection is not a rare event. Floyd et al. found that 86% of participants could recall a time having expressed manipulative affection; however, "more than half" had done so within the prior 30 days (p. 137). Primary targets of manipulative affection were friends and current or former romantic partners for relationship, target, and self-centered motives. This is considered a face-threatening act (Erbert & Floyd, 2004) and is likely a risk for receivers of affectionate messages (see Floyd, 2006).

Thus far the research reviewed has primarily focused on understanding deceptive affection in romantic relationships. Yet, deceptive affection has wide applications across contexts. To that end, researchers have recently explored deceptive affection in conjunction with health communication.

Deceptive Affection and Health Communication

Initial deceptive affection studies described how and why romantic partners utilized deceptive affection in relationships. Researchers are moving toward applying deceptive affection in new contexts, primarily focused on health communication.

Health care is a context in which providers may need to communicate support and affection, while recognizing that the job environment precludes entirely honest communication. Horan and Parker-Raley (Horan, Parker-Raley, & Cafferty 2015; Parker-Raley & Horan, 2014) described how and why emergency department (ED) staff expressed DAMs and withheld affection. Consistent with affectionate communication researchers' arguments (e.g., Floyd & Pauley, 2011), they utilized the research on emotional labor (Hochschild, 1983) to argue that deceptive affection might be employed to prevent compassion fatigue (Huynh, Alderson, & Thompson, 2008; Zeidner, Hadar, Matthews, & Roberts, 2013) and burnout (e.g., Maslach, 1982). Their findings document that health care providers do utilize DAMs and also withhold felt affection. Although their small sample sizes and self-report recall method call for additional studies of this process, it is reasonable

to argue that the study of deceptive affection has wide application. Diverse careers where emotions are at the forefront, with conceptual and theoretical connections to emotional labor provide fertile areas for research.

Researchers have also applied deceptive affection concepts to the study of safer sex and risk (Horan, 2016; Horan & Cafferty, 2017). These studies have distinguished between the communication *of* affection and communication *about* affection. For instance, having sex with someone for whom you feel love is an affectionate message. However, being deceptive about your sexual history would be deception *about* the subject of affection. Thus, deceptive affection has been recently extended to include those instances where people lie about affection. Viewing the various forms of deceptive affection together, each utilizes affection in a different way: DAMs occur when affection is the vehicle of deception, withholding affection utilizes affection as the feeling not fully expressed, and deception about affection is a form of meta-communication that entails using affection as the subject of deception (it can occur either with, or without, the expression of affection).

Horan (2016) first examined the honest or deceptive nature of individuals when communicating about their number of previous sexual partners. Extending Lucchetti's (1999) initial study on this topic, Horan found that about 60% of participants had been deceptive about their number of previous sexual partners (violating a safer sex behavior, see Planned Parenthood, 2016). Disclosure of one's number of sexual partners is one aspect of sexual history conversations, and Redlick (2017) examined multiple aspects of such conversations. She created a scale that gauged "sexual history topic avoidance," and this self-report scale had participants rate the degree to which they avoided talking about: "(1) my previous sexual experiences; (2) the number of previous sexual partners I have had; (3) my status as a virgin or non-virgin; and (4) previous relational partners whom I have been physically intimate" (p. 152). The degree to which sexual history communication was perceived as threatening was positively related to one's number of sexual partners and sexual history topic avoidance. Moreover, the perceived threat of sexual communication mediated the relationship between sexual history topic avoidance and previous number of sexual partners. Collective findings highlight the importance of studying safer sex communication and deceptive affection.

In light of the above findings, and to understand this process better, Horan (2016; Horan & Cafferty, 2017; Horan, Morgan, & Burke, 2018) has argued for the use of AET generally, and deceptive affection specifically, in studies of safer sex and risk. Still, in studying deception about sexual histories, numerous explanations have been proposed: taboo topics (Lucchetti, 1999), privacy (Nichols, 2012), and dialectical tensions (Horan, 2016; Lucchetti, 1999). From an affectionate communication perspective, Horan (2016) argued that range of tolerance for affectionate communication might explain deception about sexual histories. Potentially, discussions of safer sex globally, and sexual histories specifically, might exceed one's range of tolerance

and, therefore, individuals avoid them or act deceptively. Utilizing AET's arguments and deceptive affection research is likely telling in understanding whether, and how, people discuss sexual safety.

THEORETICAL CONSIDERATIONS

As reviewed in the opening, the study of deceptive affection is firmly rooted in the arguments of AET (Floyd, 2006). The descriptive studies of deceptive affection (Carton & Horan, 2014; Horan & Booth-Butterfield, 2013; Horan et al., 2015; Parker-Raley & Horan, 2014) not only support the theoretical arguments presented in Floyd's work, but also describe how and why affectionate feelings and messages diverge. Additionally, AET argues that affection enhances relational qualities, and as previously discussed, deceptive affection could serve as relational maintenance.

An additional AET argument is that humans have a range of tolerance for affectionate communication. As the theory explains: "the optimal range of tolerance for affection and affectionate behavior represents a range spanning the lowest sufficient amount to the highest desired amount" (Floyd, 2006, p. 174). Consequently, an individual might withhold expressing affectionate messages because it exceeds his or her range of tolerance. Likewise, an individual may express a DAM because expressing an authentic affectionate message might also exceed one's range of tolerance (Horan & Booth-Butterfield, 2013). When studying deception about affection, Horan (2016; Horan & Cafferty, 2017) argued that individuals may avoid communication about safer sex as well, because such interactions could violate their range of tolerances for affection.

The deceptive affection process can also be productively viewed through deception theories. One such theory is information manipulation theory (IMT; McCornack, 1992). This theory proposes that deception occurs when sources knowingly create messages that violate expectations of quantity, relevance, manner, and/or quality. As the theory of information manipulation most firmly focuses on the manipulation of messages/information as deception, Horan (2013a) aimed to understand how features of DAMs might account for differences in deceivers' reports of guilt, shame, and rumination. Horan had participants recall a recent DAM expressed to a romantic partner. His findings suggest that, of feelings, messages, and motives, the feelings experienced when expressing a DAM were most telling in understanding reports of guilt, shame, and rumination. McCornack has proposed an extension of his theory, IMT2, and future deceptive affection studies should consider how deceptive affection operates in conjunction with IMT2's arguments (McCornack et al., 2014).

Recently, interpersonal acceptance-rejection theory (Rohner, 1986) has been discussed in conjunction with deceptive affection (Denes, Bennett, & Winkler, 2017). In brief, the theory offers a framework for understanding the

implications of acceptance/rejection (for a review, see Rohner & Lansford, 2017). Denes et al. provide a review of the key arguments from the theory that they believe are relevant to affectionate communication. They call for research on deceptive affection to see how it operates in conjunction with acceptance/rejection. They encouraged "building on these dark-side forms of affection to further theorize about the gray area between acceptance and rejection" (p. 500).

A final theory that might have research application within the deceptive affection process is predicted outcome value theory (POV; Sunnafrank, 1986). This theory was originally designed to understand relationship development. That is, the theory argues that individuals desire profitable relationships and, upon initial interactions, generate predictions regarding costs and rewards associated with a potential relationship with a new other. Theoretically, if a profitable relationship is anticipated, communication will occur in such a way to encourage relationship development. Alternatively, if a costly relationship is anticipated, communication will occur in a way to discourage relationship development. Possibly, DAMs may be enacted in order to enhance the POV of a specific initial interaction. The theory has been recently extended to understanding POV judgments in ongoing relationships (Ramirez, Sunnafrank, & Goei, 2010). AET and transgressions have been previously studied together (Horan, 2012), and some consider deception to be a transgression (Mettes & Cupach, 2007). Consequently, the use and/or discovery of deception is considered an unexpected transgression that could result in the re-evaluation of POV in an established relationship. Future studies should explore this idea.

FUTURE RESEARCH

This review has discussed research describing deceptive affection—still, much remains to be learned about this process. Therefore, future research would be wise to consider the following. First, replication of deceptive affection diary studies (Carton & Horan, 2014; Horan and Booth-Butterfield, 2013) in same-sex relationships would extend our understanding of this process. Potentially, society's negative views of same-sex relationships would limit the freedoms of same-sex couples to engage in public displays of affection, resulting in higher levels of withheld affection (Carton & Horan, 2014). Second, studies could move from general deceptive affection to understanding specific messages or enactments. Such studies should incorporate display rules to understand more fully how individuals employ deceptive affection. Example messages to study include deception about aspects of sexual history conversations (Horan, 2016; Redlick, 2017) and pretending to orgasm (Denes, Horan, & Bennett, 2018).

Third, it is important to move from recognition that a concept or process exists, to the outcomes it generates. It seems evident that DAMs cause little

distress in communicators, but that claim is based on a single experimental manipulation. Future studies should adopt different experimental designs to examine the risks of communicating DAMs or withholding affection. Additionally, there may be numerous effects of DAMs including relationship satisfaction indicators and recognition by receivers of deceptive affection. Potentially, might deceptive affection be met with a deceptive affection response? Using dyadic methods would offer an understanding of whether deceptive affection turns into a cycle or pattern in relationships.

Fourth, some participants in deceptive affection studies have indicated that they do not lie and, therefore, were excluded from analyses. For example, 19.08% of Horan's (2013b) participants reported they never used deceptive affection, and three of the 57 diaries submitted in Horan and Booth-Butterfield's (2013) diary study reported zero instances of DAMs. In follow-up surveys, Horan and Booth-Butterfield's (2013) participants indicated that they did not communicate deception to their partners. Individuals who claim to never lie should be studied to further understand source characteristics of the "always honest," interviewed to understand how they manage their honesty, reactions to their blatant honesty, and studied to understand whether their complete commitment to honest communication is myth or reality. As Horan speculated: "It is unknown if communicators truly never communicate deceptive affection or, more probable, if they do not view it as deception…Since deceptive affection may [be] viewed as potentially altruistic or as not imposing harm to their partner, actors may simply not view this as deception and, accordingly, indicated they never engage in this behavior" (p. 19).

Fifth, as guidance for general affectionate communication research, deceptive affection researchers have encouraged future affectionate communication studies to include measures that distinguish between honest and deceptive affection (Horan & Booth-Butterfield, 2011). Doing so will paint a more complete picture of affectionate communication and AET.

Sixth, an area for future research pertains to measurement development. Quantitative studies of deceptive affection have involved the use of single item measures (Gillen & Horan, 2013; Horan 2013b) as well Trask et al.'s scale. That said, authors of this chapter have concerns over the conceptual/operational fit with Trask et al.'s deceptive affection scale. Thus, future research should consider the best ways in which to measure the frequency of deceptive affection. Researchers are encouraged to base scale items on initial diary studies of deceptive affection. Using the findings of Horan and Booth-Butterfield (2013) and Carton and Horan (2014) to guide scale development should enhance the validity of item development.

Finally, understanding whether deceptive affection is a dark side issue warrants exploration. The question of whether deceptive affection specifically, and deception generally, is a dark side process is broad. Note that recent volumes addressing dark side issues do not have general deception chapters

(Cupach & Spitzberg, 2011; Spitzberg & Cupach, 2007b), yet they do address the content in some chapters. Still, in other volumes addressing dark side issues, there is an affection chapter (Floyd & Pauley, 2011) and a chapter on avoidance and secrets (Afifi, Caughlin, & Afifi, 2007). As acknowledged in previous dark side chapters, there is variance in how messages operate. That is, Spitzberg and Cupach (2007a), when discussing social support as a research example, speculate from a dark side perspective: "Under what conditions does social support promote health, and under what conditions does social support impair health?" (p. 7). Similarly reasoned here, the question becomes under what conditions might deceptive affection result in positive outcomes for relational and health communication and how might it be problematic? Though reporters and students have regularly asked us, "Is deceptive affection good or bad?", the research reported highlights that a simplistic answer is not possible. When asked, is this a dark side of communication process, our answer then is: perhaps. Deceptive affection likely produces both productive and unproductive outcomes; future studies should examine such outcomes.

CONCLUSION

Years of deception research, as discussed in various chapters in this volume, have documented the frequent nature of deceptive communication. Concurrently, the programmatic theoretical studies of affectionate communication have revealed the numerous benefits associated with affectionate messages (see Floyd, 2006). Though perhaps a counterintuitive notion, combining the study of deception and affection into deceptive affection has resulted in a body of research that has only begun to explore this area. Though the research reviewed here offered a descriptive understanding of deceptive affection, and highlighted potential risks and benefits, theoretical explanations, and emerging trends, much remains to be learned about the communication and implications of deceptive affection across a variety of relationships in various contexts. The discussion presented here represents just the beginning of what will, hopefully, be years of research in this area. Ultimately, deceptive affection is not a rare form communication, and therefore, people will continue to use angry hugs and withhold love. Time will only tell, through years of study, the true implications of deceptive affection.

REFERENCES

Afifi, T., Caughlin, J., & Afifi, W. (2007). The dark side (and light side) of avoidance and secrets. In B. H. Spitzberg & W. R. Cupach (Eds.), *The dark side of interpersonal communication* (pp. 61–92). Mahwah, NJ: Lawrence Earlbaum.

Andersen, P. A., & Guerrero, L. K. (1998). Principles of communication and emotion in social interaction. In P. A. Andersen & L. K. Guerrero (Eds.), *Handbook of communication and emotion: Research, theory, applications, and contexts* (pp. 49–96). San Diego, CA: Academic Press.

Buller, D. B., & Burgoon, J. K. (1998). Emotional expression in the deception process. In P. A. Andersen & L. K. Guerrero (Eds.), *Handbook of communication and emotion: Research, theory, applications, and contexts* (pp. 381–398). San Diego, CA: Academic.

Carton, S. T., & Horan, S. M. (2014). A diary examination of romantic and sexual partners withholding affectionate messages. *Journal of Social and Personal Relationships, 31,* 221–246. https://doi.org/10.1177/0265407513490587.

Coward, N. www.goodreads.com/quotes/178186. Retrieved February 6, 2018.

Cupach, W. R., & Spitzberg, B. H. (2011). *The dark side of close relationships II.* New York: Routledge.

Dainton, M., & Gross, J. (2008). The use of negative behaviors to maintain relationships. *Communication Research Reports, 25,* 179–191. https://doi.org/10.1080/08824090802237600.

Denes, A., Bennett, M., & Winkler, K. L. (2017). Exploring the benefits of affectionate communication: Implications for interpersonal acceptance-rejection theory. *Journal of Family Theory & Review, 9,* 491–506. https://doi.org/10.1111/jftr.12218.

Denes, A., Horan, S. M., & Bennett, M. (2018). *"Faking it" as deceptive affection: Exploring the authenticity of orgasm and relational quality indicators.* Paper presented at the annual meeting of the National Communication Association, Salt Lake City, UT.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74,* 63–79. https://doi.org/10.1037/0022-3514.74.1.63.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995. https://doi.org/10.1037/0022-3514.70.5.979.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues.* Englewood Cliffs, NJ: Prentice Hall.

Erbert, L. A., & Floyd, K. (2004). Affectionate expressions as face-threatening acts: Receiver assessments. *Communication Studies, 55,* 254–270. https://doi.org/10.1080/10510970409388618.

Floyd, K. (1997a). Communicating affection in dyadic relationships: An assessment of behavior and expectancies. *Communication Quarterly, 45,* 68–80. https://doi.org/10.1080/01463379709370045.

Floyd, K. (1997b). Knowing when to say "I love you": An expectancy approach to affectionate communication. *Communication Research Reports, 14,* 321–330. https://doi.org/10.1080/08824099709388675.

Floyd, K. (2001). Human affection exchange: I. Reproductive probability as a predictor of men's affection with their sons. *Journal of Men's Studies, 10,* 39–50. https://doi.org/10.1080/08824099709388675.

Floyd, K. (2006). *Communicating affection: Interpersonal behavior and social context.* Cambridge, UK: Cambridge University Press.

Floyd, K., & Pauley, P. M. (2011). Affectionate communication is good, except when it isn't: On the dark side of expressing affection. In W. R. Cupach & B. H. Spitzberg (Eds.), *The dark side of close relationships II* (pp. 145–173). New York, NY: Routledge.

Floyd, K., & Voloudakis, M. (1999). Affectionate behavior in adult platonic friendships. *Human Communication Research, 25,* 341–369. https://doi.org/10.1111/j.1468-2958.1999.tb00449.x.

Gillen, H., & Horan, S. M. (2013). Toward an understanding of the relationships among deceptive affection, deceptive beliefs, and relational qualities. *Communication Research Reports, 30,* 352–358. https://doi.org/10.1080/08824096.2013.836629.

Guerrero, L. K., & Bachman, G. F. (2006). Associations among relational maintenance behaviors, attachment-style categories, and attachment dimensions. *Communication Studies, 57,* 341–361. https://doi.org/10.1080/10510970600845982.

Guthrie, J., & Kunkel, A. (2013). Tell me sweet (and not-so-sweet) little lies: Deception in romantic relationships. *Communication Studies, 64,* 141–147. https://doi.org/10.1080/10510974.2012.755637.

Hochschild, A. R. (1983). *The managed heart.* Berkeley, CA: University of California Press.

Horan, S. M. (2012). Affection exchange theory and perceptions of relational transgressions. *Western Journal of Communication, 76,* 109–126. https://doi.org/10.1080/10570314.2011.651548.

Horan, S. M. (2013a, November). *Do message features help explain communicators' reactions to expressed deceptive affectionate messages?* Paper Presented at the Annual Meeting of the National Communication Association, Washington, DC.

Horan, S. M. (2013b, April). *Further understanding affectionate communication as maintenance*. Paper Presented at the Annual Meeting of the Eastern Communication Association, Pittsburgh, PA.

Horan, S. M. (2016). Further understanding sexual communication: Honesty, deception, safety, and risk. *Journal of Social and Personal Relationships, 33,* 449–468. https://doi.org/10.1177/0265407515578821.

Horan, S. M., & Booth-Butterfield, M. (2010). Investing in affection: An investigation of affection exchange theory and relational qualities. *Communication Quarterly, 58,* 394–413. https://doi.org/10.1080/01463373.2010.524876.

Horan, S. M., & Booth-Butterfield, M. (2011). Is it worth lying for? Physiological and emotional implications of recalled deceptive affection. *Human Communication Research, 37,* 78–106. https://doi.org/10.1111/j.1468-2958.2010.01394.x.

Horan, S. M., & Booth-Butterfield, M. (2013). Understanding the routine expression of deceptive affection in romantic relationships. *Communication Quarterly, 61,* 195–216. https://doi.org/10.1080/01463373.2012.751435.

Horan, S. M., & Cafferty, L. A. (2017). Condom communication: Reports of sexually active young adults' recent messages with new partners. *Journal of Health Communication, 22,* 763–771. https://doi.org/10.1080/10810730.2017.1355420.

Horan, S. M., Morgan, T., & Burke, T. J. (2018). Sex and risk: Parental messages and associated safety/risk behavior of adult children. *Communication Quarterly, 66*(4), 403–422. https://doi.org/10.1080/01463373.2017.1418404.

Horan, S. M., Parker-Raley, J., & Cafferty, L. A. (2015, November). *Emotional dissonance: Withholding affectionate messages in emergency departments*. Paper Presented at the Annual Meeting of the National Communication Association, Las Vegas, NV.

Huynh, T., Alderson, M., & Thompson, M. (2008). Emotional labour underlying care: An evolutionary concept analysis. *Journal of Advanced Nursing, 64,* 195–208. https://doi.org/10.1111/j.1365-2648.2008.04780.x.

Kaplar, M. E., & Gordon, A. E. (2004). The enigma of altruistic lying: Perspective differences in what motivates and justifies lie telling within romantic relationships. *Personal Relationships, 11,* 489–507. https://doi.org/10.1111/j.1475-6811.2004.00094.x.

Levine, T., McCornack, S., & Avery, P. (1992). Sex differences in emotional reaction to detected deception. *Communication Quarterly, 40,* 289–296. https://doi.org/10.1080/01463379209369843.

Lucchetti, A. E. (1999). Deception in disclosing one's sexual history: Safe-sex avoidance or ignorance? *Communication Quarterly, 47,* 300–314. https://doi.org/10.1080/01463379909385561.

Maslach, C. (1982). *Burnout: The cost of caring.* Englewood Cliffs, NJ: Prentice-Hall.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs, 59,* 1–16. https://doi.org/10.1080/03637759209376245.

McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology, 33,* 348–377. https://doi.org/10.1177/0261927X14534656.

Metts, B. H., & Cupach, W. R. (2007). Responses to relational transgressions: Hurt, anger, and sometimes forgiveness. In B. H. Spitzberg & W. R. Cupach (Eds.), *The dark side of interpersonal communication* (pp. 243–274). Mahwah, NJ: Lawrence Earlbaum.

Morman, M. T., & Floyd, K. (1998). "I love you man": Overt expressions of affection male-male interaction. *Sex Roles, 38,* 872–881. https://doi.org/10.1023/A:1018885417249.

Nichols, W. L. (2012). Deception versus privacy management in discussions of sexual history. *Atlantic Journal of Communication, 20,* 101–115. https://doi.org/10.1080/15456870.2012.665346.

Parker-Raley, J., & Horan, S. M. (2014, November). *How and why emergency department staff express inauthentic affectionate messages.* Paper Presented at the Annual Meeting of the National Communication Association, Chicago, IL.

Pennebaker, J. W., & Beall, S. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology, 95,* 274–281. https://doi.org/10.1037/0021-843X.95.3.274.

Planned Parenthood (2016). *Talking with your partner about safe sex.* Retrieved January 29, 2016 from https://www.plannedparenthood.org/planned-parenthood-southwestern-oregon/oregon-teens/talking-your-partner-about-safe-sex.

Ramirez, A., Sunnafrank, M., & Geoi, A. (2010). Predicted outcome value theory in ongoing relationships. *Communication Monographs, 77,* 27–50. https://doi.org/10.1080/03637750903514276.

Redlick, M. (2017). The perceived threat of sexual communication, number of previous sexual partners and topic avoidance in romantic relationships. *Psychology & Sexuality, 8,* 148–157. https://doi.org/10.1080/19419899.2017.1316768.

Rogers, T., Zeckhauser, R., Gino, F., Norton, M., & Schweitzer, M. (2017). Artful Paltering: The risks and rewards of using truthful statements to mislead others. *Journal of Personality and Social Psychology, 112,* 456–473. https://doi.org/10.1037/pspi0000081.supp.

Rohner, R. P. (1986). *The warmth dimension: Foundations of parental acceptance–rejection theory.* Beverly Hills, CA: Sage.

Rohner, R. P., & Lansford, J. E. (2017). Deep structure of the human affectional system: Introduction to interpersonal acceptance-rejection theory. *Journal of Family Theory & Review, 9,* 426–440. https://doi.org/10.1111/jftr.12219.

Scholl, J. C., & O'Hair, D. (2005). Uncovering beliefs about deceptive communication. *Communication Quarterly, 53,* 377–399. https://doi.org/10.1080/01463370500101352.

Spitzberg, B. H., & Cupach, W. R. (2007a). Disentangling the dark side of interpersonal communication. In W. R. Cupach & B. H. Spitzberg (Eds.), *The dark side of interpersonal communication* (2nd ed., pp. 3–28). Mahwah, NJ: Lawrence Earlbaum.

Spitzberg, B. H., & Cupach, W. R. (2007b). *The dark side of interpersonal communication* (2nd ed.). Mahwah, NJ: Lawrence Earlbaum.

Sunnafrank, M. (1986). Predicted outcome value during initial interactions: A reformulation of Uncertainty Reduction Theory. *Human Communication Research, 13,* 3–33. https://doi.org/10.1111/j.1468-2958.1986.tb00092.x.

Trask, S. L., Horstman, H. K., & Hesse, C. (2016). *Deceptive affection across relational contexts: A group comparison of romantic relationships, cross-sex friendships, and friends with benefits.* Presented at National Communication Association Annual Conference, Philadelphia, PA.

Zeidner, M., Hadar, D., Matthews, G., & Roberts, R. (2013). Personal factors related to compassion fatigue in health professionals. *Anxiety Stress and Coping, 26,* 595–609. https://doi.org/10.1080/10615806.2013.777045.

Deceiving for and During Sex

Gayle Brewer

Romantic and sexual relationships form an important part of the social land-scape. They are however, as with all other social relationships, vulnerable to deception. The current chapter outlines the use of deception to obtain sex (i.e., false advertising) and during sex (i.e., pretending to experience orgasm and infidelity), with particular focus on important differences between men and women.

Deception to Obtain Sex

Parental Investment and Sex Differences

The selection, attraction, and retention of a suitable partner are of fundamental importance. It is perhaps not surprising then that deception is frequently employed to gain advantage (e.g., to appear more attractive and attract a partner that would normally be unobtainable). The type of partner or relationship sought and the consequences of poor decision-making are however markedly different for men and women, which as a result influence the level or type of deception employed. Differences reflect the minimum levels of investment required by each sex to produce a healthy child.

Though only one gamete is required for successful conception, sperm are relatively 'cheap' to produce and men produce millions of sperm per ejaculate. After conception, men may choose to support women during pregnancy and care for the child after birth or they may choose to make no further investment in the woman or his child. Hence, the *minimum* investment

G. Brewer (✉)
University of Liverpool, Liverpool, UK
e-mail: gbrewer@liverpool.ac.uk

required of men (energy, time) is relatively low. Furthermore, while men's age may exert some impact on the health of the child (Jenkins, Aston, Pflueger, Cairns, & Carrell, 2014), men are capable of producing children from puberty to late adulthood. Men are, therefore, physically capable of producing a large number of children. As men may conceive a child during a sexual encounter and then immediately search for another partner, male reproductive output (i.e., the number of children produced) is to a large degree restricted by his access to willing fertile partners rather than his capacity to conceive them.

In contrast to the aforementioned paternal investment, the minimum maternal investment is much greater. Women produce only one gamete (ovum) per month, and likelihood of conceptualization varies across the monthly menstrual cycle. If conception does occur, women experience a nine-month energy-intensive pregnancy, followed by an extended period of lactation. Hence, one sexual encounter may result in substantial investment for a number of years. Furthermore, female fertility is restricted across the life span. Though women are able to reproduce after puberty, where women exert control over their own reproduction conception is typically delayed. Such delay reflects the risks posed during childbirth to younger mothers and their children (Hendrie, Brewer, Lewis, & Mills, 2014). Menopause prevents conception in later life though women typically end reproduction at an earlier age, reflecting the complications that arise for older women and her children. Hence, women's reproductive output is restricted and behavior intended to recruit support from partners and increase the well-being of each child is more beneficial than behavior attempting to increase the number of partners or children.

Together, the biological pressures, minimum investment in each child, and potential reproductive output experienced by men and women suggest that each sex should differ with regard to the relationship and partner they prefer. Consequently, each sex should differ with regard to the deception they use in order to obtain sex. The theory of parental investment (Trivers, 1972) asserts that the sex which invests the least in each offspring (i.e., men) should favor short-term relationships while the sex with the greatest investment (i.e., women) should prefer long-term committed relationships. This strategy provides men with access to partners able to raise his children and allows women to secure the investment and support which lessens the burden of childcare. In one seminal study investigating relationship preferences, men and women were approached on a college campus and asked (a) Would you go out with me tonight? (b) Would you come over to my apartment tonight? or (c) Would you go to bed with me tonight? Results indicated that approximately half of those questioned (regardless of sex) would agree to the date. Few women would agree to go to the man's apartment and no women would agree to sexual intercourse. In contrast, men were more likely to agree to go to a woman's apartment than to agree to a date and were most likely to

agree to sexual intercourse, with approximately three-quarters of men willing to accept the offer (Clark & Hatfield, 1989). A similar pattern of results has been reported by subsequent studies, despite lower levels of acceptance. For example, Hald and Høgh-Olesen (2010) report that similar numbers of women and men agreed to go on a date. When asked to visit the prospective partner's apartment 22% of men and 8% of women agreed while when presented with a sexual invitation ("Would you go to bed with me?") 38% of men and 2% of women agreed. Indeed, there is a substantial amount of research indicating that in general, men are more sexually promiscuous and women are more cautious regarding sexual behavior (Fletcher, Kerr, Li, & Valentine, 2014; Schmitt et al. 2003).

Researchers have also established sex differences are also apparent for the type of partner preferred. In particular, men, whose reproductive output is limited by access to fertile women, place a greater importance on the physical attractiveness of a mate (Buss & Schmitt, 1993; Kamble, Shackelford, Pham, & Buss, 2014). Those physical features contributing to assessments of physical attractiveness are not arbitrary and reveal important information about the age and fertility of a woman. For example, body shape (and the "hour glass" figure) is commonly discussed with reference to female attractiveness. Prior to puberty, girls and boys display a similar waist-to-hip ratio. At puberty, estrogen and testosterone stimulate the accumulation of fat in sex-specific areas. Therefore, women typically display a lower waist-to-hip ratio (approximately 0.67–0.80) than men (approximately 0.85–0.95). Female ratios do however increase after childbirth and menopause. Waist-to-hip ratio is also associated with a range of health conditions such as cardiovascular disease and breast cancer, together with likelihood of conception. Hence, waist-to-hip ratio provides important information about women's age and reproductive status (see Singh, 2002 for a review). Other physical attributes such as levels of body fat (i.e., body mass index) also influence ratings of physical attractiveness and perceived health, although these characteristics may signal different information. For example, body mass index may reveal the ability to endure pregnancy and energy-intensive lactation, while waist-to-hip ratio indicates youth and fertility.

In contrast, women, whose reproductive output is restricted by the time and energy required to sustain pregnancy, lactation, and child-rearing, place comparatively greater emphasis on the resources held by a mate and their ability to acquire resources (Fales et al., 2016). These resources reduce her vulnerability and increase the likelihood that her children will survive, particularly in harsh environments. Hence, women report a preference for partners that are older, ambitious, and hardworking (Buss & Schmitt, 1993). As men may possess resources but choose not to invest these in a partner or subsequent children, the extent to which he is committed to a woman and her children is also important (Buss & Schmitt, 1993). Unsurprisingly, it appears that each sex is aware of the traits and type of relationship preferred

by the opposite sex. Furthermore, these preferences influence the manner in which men and women advertise (honestly or otherwise) for romantic and sexual partners.

Advertisement and False Signalling

Each sex advertises a range of qualities desired by potential partners, whether consciously or unconsciously. Furthermore, these qualities appear to differ for men and women. For example, dating advertisements posted by men are more likely to describe resources (e.g., possession of a professional job) and seek physical attractiveness (e.g., nice body) in a partner, whereas women display the opposite pattern (i.e., describe their own physical attractiveness and seek resources; Alterovitz & Mendelsohn, 2009; Waynforth & Dunbar, 1995). To increase the number or quality of potential mates attracted by the advertisement, men and women may emphasize, exaggerate, or artificially create positive characteristics when advertising themselves to potential partners in order to create a positive impression (Toma & Hancock, 2010). Consistent with the aforementioned differences in partner and relationship preferences, men and women focus on different traits when engaging in deceptive self-presentation. For example, women may conceal or alter their age, which is related to female fertility and physical attractiveness, in order to increase their access to high-quality partners (Pawlowski & Dunbar, 1999). Indeed, women are more likely than men to report that a previous partner has exaggerated their income or ambitions, or has exaggerated their feelings in order to obtain sexual access. In contrast, men are more likely than women to report that they have previously been sexually led on (Haselton, Buss, Oubaid, & Angleitner, 2005).

False Advertisement Online

A substantial and increasing number of people advertise online in order to meet social, romantic, or sexual partners (Pew Research Center, 2016). Sites may be focused on the formation of either committed or casual relationships and are often targeted at a particular market (e.g., professionals, single parents, seniors) or shared interests (e.g., sporting activity, political affiliation). Users provide a range of background information (e.g., demographics, personality, physical traits) together with their partner and relationship preferences. Daters typically also display a personal photograph. The nature of online interactions provides ample opportunities for deception. Indeed, daters appear aware of this and deception or misrepresentation are perceived disadvantages of online dating (Brym & Lenton, 2001; Rochadiat, Tong, & Novak, 2017). Users may deceive others about the most basic personal information; for example, 'gender switching' is one of the most common forms of deception. Whitty (2002) found that 18% of men and 11% of women using

online chat rooms reported having lied about their sex. Online users rate this kind of deception as the most distressing and is particularly of concern for those searching for romantic compared to social chat partners (Stieger, Eichinger, & Honeder, 2009). The distress caused by gender switching may reflect the level of time and other resources that individuals invest into the development of a relationship which ultimately has no future. Alternatively, deceived individuals may experience stigma or a loss of social status if the deception is revealed to other people.

Of course, deception may impact relationship development, as well as its success. For example, if a man or woman meets their online partner in person and finds them to be significantly different from their profile, they may terminate the relationship. Hence, online dating-based deception may be more subtle. Users may also adopt a strategic approach to deception, altering the information they present (i.e., false advertising) in response to the parameters of the online dating site. For example, users may change their age to ensure that they are not excluded from online searches (Ellison, Heino, & Gibbs, 2006), and it is difficult for users to identify this form of deception. Research indicates that deception within dating profiles is associated with specific linguistic cues (e.g., singular pronouns, negations), however while computer programmes can detect these cues, human judges are unable to (Toma & Hancock, 2012). It is important to note that online daters often report that their own profiles are accurate (Hancock & Toma, 2009), which may reflect the tendency to create online profiles that reflect their 'ideal self' rather than their actual self (Ellison, Heino, & Gibbs, 2006). Therefore, they may not necessarily view this practice as deceptive.

Individuals who use online dating applications typically present themselves in a manner that appeals to potential partners (Sedgewick, Flath, & Elias, 2017; Ward, 2017). This may include highlighting desirable qualities, concealing undesirable material, or falsely reporting the presence of valued physical or non-physical traits. Men and women appear to enhance different traits, which are targeted at the qualities most sought by potential partners. For example, men are more likely to enhance status whereas women are more likely to emphasize physical appearance (Hitsch et al., 2009). This may be an effective strategy for those wishing to increase their online popularity and the number of potential partners available. Indeed, physical appearance and income predict the number of responses women and men receive via online dating sites (Hitsch et al., 2009). Appearance-based deception is also influenced by the traits sought by potential partners. For example, when men and women misrepresent their physical appearance, women are more likely to lie about their weight, whereas men are more likely to lie about height (Toma, Hancock, & Ellison, 2008). Indeed, independent raters report that approximately one-third of profile photographs are not accurate, with female photographs viewed as less accurate than male photographs (Hancock & Toma, 2009). Perhaps as a consequence, many online daters believe that people

misrepresent their physical appearance (Gibbs, Ellison, & Heino, 2006) and are most likely to doubt the authenticity of attractive versus unattractive profile photographs (Lo, Hsieh, & Chiu, 2013).

Detecting Deception

Both men and women are adapted to both display and detect deception, though important sex differences may occur. In particular, women may be more sensitive than men to the aforementioned dishonest advertising, especially when there are important reproductive consequences (Dimoulas, Wender, Keenan, Gallup, & Goulet, 1998). Consistent with expectations, women report greater distress in response to deception than men (Docan-Morgan & Docan, 2007; Haselton et al., 2005) and expect the opposite sex to engage in deception (Keenan, Gallup, Goulet, & Kularni, 1997). In particular, compared to women in committed romantic relationships, single women are better at detecting men "faking good" i.e., men who use enhanced descriptions of themselves in a manner that might appeal to potential partners (Johnson et al., 2004).

This form of self-enhancing deception may be more frequently employed by men seeking short-term relationships (Keenan, Gallup, & Falk, 2003). Women detecting this deception may reduce the likelihood that they are left "holding the baby" without the support of a suitable long-term partner. Women may also avoid the reputational damage associated with short-term relationships. Of course, the likelihood that women will conceive varies across the menstrual cycle. Therefore, the consequences of deception also vary. Women who do not detect false promises of commitment at low levels of fertility may suffer distress or reputational damage; at high levels of fertility women may conceive a child and unwillingly face energy-intensive pregnancy and child-rearing alone. In consequence, it is not just women's preferences and sexual behavior which varies across the menstrual cycle; it is also her responses to deception. Commitment skepticism is higher among naturally cycling women (i.e., those who do not use hormonal contraceptives) during fertile phases of the menstrual cycle (Peterson, Carmen, & Gehr, 2013).

Sex differences in response to the nature of the deception also reflect sex-specific evolutionary pressures and partner preference. Haselton et al. (2005) report that women, more so than men, believe that they would be distressed to learn that a romantic partner had exaggerated their income or status, that they had exaggerated their feelings in order to obtain sex, or that a romantic interest had concealed a current relationship with another person. Women are also more likely than men to be distressed by a partner's lack of interest in a committed romantic relationship following sexual activity (Haselton et al., 2005). In contrast, men are more likely than women to be distressed if a partner falsely suggests that they are willing to engage in sexual

intercourse (Haselton et al., 2005). Reactions to deception are of course influenced by the type of relationship people wish to develop. For example, men and women who are oriented toward short-term relationships are less distressed by commitment deception (i.e., pretending to be more committed than he or she is) than those focused on long-term relationships (Haselton et al., 2005).

Sexual History-Based Deception

Those attempting to persuade a person to engage in sexual activity, whether part of an established relationship or not, may attempt to conceal their sexual history. Full disclosure of previous sexual history may threaten the attractiveness of a person and the development of the relationship if their sexual history is perceived negatively. In particular, disclosure may lead to rejection, social stigma, or embarrassment. Hence, many men and women avoid revealing this information. For example, Lucchetti (1999) reports that one-third of sexually active students have avoided disclosing their sexual history to at least one partner prior to sexual involvement, while Stebleton and Rothenberger (1993) reveal that all men participating in their research had failed to disclose their sexual history to at least one partner. In addition to deceiving partners about the number of previous partners, deception may also occur in relation to the identity of previous partners or previous incidents of infidelity (Williams, 2001). Those who have been unfaithful in previous relationships are more likely to be unfaithful in their current relationship (Adamopoulou, 2013), hence concealing prior infidelity may reduce the suspicion of the current partner. In a similar manner, concealing the identity of previous partners may reduce the likelihood that partners would seek to limit contact with previous mates.

Partners who fail to discuss their prior sexual relationships may experience less intimacy and emotional closeness with current partners. Avoiding the subject may, therefore, be problematic. Hence, men and women may choose to actively lie about their sexual history rather than refuse to discuss the subject (i.e., lie by omission). Approximately one-third of those questioned by researchers confess to deceiving partners by lowering the number of previous sexual relationships (Knox, Schacht, Holt, & Turner, 1993), while approximately half of those questioned plan to lower the number of previous partners disclosed to future relationship partners (Cochran & Mays, 1990). The impact of these deceptions may extend beyond the quality of the relationship itself to important health consequences. Desiderato and Crawford (1995) report that 42% of those who had currently or previously had an STD did not tell partners before sexual involvement, and 17% of those who were diagnosed as HIV positive did not inform partners about their status. Therefore, those engaging in unprotected sexual activity may be at substantial risk of infection from deceptive partners.

DECEPTION DURING SEXUAL ACTIVITY

Pretending Orgasm

Researchers often focus on the use of deception by men and women wishing to attract a partner. Of course, deception may also occur during the sexual act itself. In particular, though 75–90% of women do not consistently orgasm during sexual activity (Bancroft, Loftus, & Long, 2003) and a substantial minority (5–10%) do not experience orgasm at all (Lloyd, 2005), women often pretend to have experienced orgasm during sexual intercourse. A range of cues (e.g., vocalizations, breathing rate, body movements) may be used to falsely indicate that orgasm has occurred. This form of deception is relatively common. Indeed, Brewer and Hendrie (2011) report that 56% of women vocalize (e.g., scream, moan) when they are not going to orgasm over 70% of the time, and 70% of women do so over 50% of the time.

With regard to the motivation for this deception, research suggests that women frequently pretend to experience orgasm in order to enhance the relationship or reassure their partner. For example, 87% of women report using vocalizations to boost their partner's self-esteem (Brewer & Hendrie, 2011), while 70% report pretending in order to avoid partner distress (Muehlenhard & Shippee, 2010). Hence, pretending to experience orgasm may provide the impression of sexual and overall relationship satisfaction. This appears to be a successful strategy as men whose partners appear to frequently orgasm report higher relationship satisfaction (Kaighobadi et al., 2012). In this manner, falsely pretending to experience orgasm—if the deception is undetected—may strengthen the relationship.

Furthermore, women may pretend to experience orgasm in order to reduce the likelihood that the partner will be unfaithful or end the relationship (Muehlenhard & Shippee, 2010). Hence, those pretending to orgasm are also more likely to engage in behaviors intended to enhance the relationship. McCoy et al. (2014) found that those who pretend to experience orgasm in order to improve their partner's sexual experience are more likely to engage in a range of mate retention behaviors (i.e., actions intended to maintain the existing relationship). These include direct mate guarding such as keeping a partner under surveillance (e.g., "Insisted that my partner spend all their free time with me"); intrasexual negative inducements such as acts intended to threaten or manipulate potential rivals (e.g., "Yelled at other men who looked at her"); positive inducements such as providing gifts, favors, or affection (e.g., "Went out of my way to be kind, nice, and caring"); public signals of possession including acts intended to signal "possession" of a partner (e.g., "Bragged about him to other women"); and intersexual negative inducements, referring to acts intended to threaten or manipulate a partner to remain faithful (e.g., "Became jealous when he went out without her"). Furthermore, McCoy et al. (2014) discovered that pretending to experience orgasm in order to hide one's sexual disinterest

is related to all forms of mate retention except the public signals of possession. While women often intend to strengthen their romantic relationships (Kaighobadi et al., 2012), if their partner detects a faked orgasm, relational trust and intimacy may decline, producing negative outcomes for the relationship itself.

Though women are more likely than men to pretend to experience orgasm during sexual intercourse (Thornhill, Gangestad, & Comer, 1995), men also engage in this form of deception. Specifically, 25% of heterosexual men report having pretended to experience orgasm on at least one occasion (Muehlenhard & Shippee, 2010). This may involve a range of verbal signals (e.g., moaning, saying that they were close to orgasm) or changing body movement (e.g., increasing strength or speed of thrusting prior to "finishing"). Compared to women, men are more likely to change body movements (e.g., thrusts) and less likely to moan or alter their breathing rate (Muehlenhard & Shippee, 2010). Similar to women, men may pretend to experience orgasm because they feel that orgasm has taken too long and want the sex to end or wish to avoid hurting their partner's feelings. They are also more likely to pretend while intoxicated or after experiencing orgasm earlier the same day (Muehlenhard & Shippee, 2010). For example, Muehlenhard and Shippee (2010) report that one man explained "One night after a couple hours of heavy drinking I was talking to this girl on my floor and apparently I was hitting on her. One thing led to another and I started sobering up during sex so I faked to make her go away… She is unattractive ⁄ annoying [and I] wanted to get her off me…when my senses came about and I took my drunk goggles off" (p. 558). Men may also choose to pretend when they do not want the partner to know that they have not experienced orgasm, which may reflect societal expectations that men should always be prepared for intercourse (Zilbergeld, 1999). For a minority of men, pretending may be used to conceal incidence of premature ejaculation (Steiner, 1981). However, relatively few studies have investigated the use of pretending to experience orgasm by men, and additional research in this area is required.

Factors Influencing the Likelihood of Pretending to Experience Orgasm

There is considerable variation with regard to the frequency of pretending to experience orgasms and motivations for this behavior. The status and quality of the sexual relationship itself are particularly important. For example, pretending to experience orgasm is most common among single compared to married women (Darling & Davidson, 1986), and more frequent among lesbian and bisexual women compared to heterosexual women (Cooper, Conner, & Fauber, 2010). Furthermore, women whose partners are less tolerant and agreeable are most likely to pretend to experience orgasm (Ellsworth & Bailey, 2013). Numerous studies have highlighted the relationship between pretending to experience orgasm and the risk of infidelity. For example,

Kaighobadi et al. (2012) report that women at greater perceived risk of a partner's infidelity are most likely to pretend to experience orgasm. Furthermore, women with a greater number of lifetime sexual partners (Wiederman, 1997) or those who have been or are likely to be unfaithful (Ellsworth & Bailey, 2013) are more likely to engage in this behavior.

Attitudes toward sexual behavior or romantic relationships (not necessarily limited to the current partner) also influence this form of deception. For example, those engaging in sexual behavior for reasons associated with insecurity, or attainment of a goal and emotional or physical reasons, are most likely to pretend to experience orgasm (McCoy et al., 2014). Furthermore, reasons for engaging in sexual behavior also influence specific motivations for pretending to experience orgasm. Those pretending to experience orgasm in order to improve their partner's sexual experience are most likely to be sexually active for insecurity, physical, or emotional reasons, while pretending to experience orgasm in order to hide sexual disinterest is associated with sexual activity for insecurity, physical reasons, and goal attainment. Pretending for deception and manipulation is associated with insecurity, emotional and physical reasons, and goal attainment motivations for sexual behavior (McCoy et al., 2014).

In addition, a range of individual differences that are relatively constant over time influence the frequency with which women pretend to experience orgasm or their motivation for doing so. For example, Machiavellianism, characterized by a manipulative interpersonal style and willingness to exploit others, is associated with a range of sexual and relationship behaviors. Previous research reports that women with high levels of Machiavellianism are more likely to pretend to experience orgasm in order to deceive and manipulate their partner (Brewer, Abell, & Lyons, 2016). Highlighting the complexity of the subject and the manner in which the characteristics of a relationship may interact with individual differences, the association between Machiavellianism and pretending orgasm was moderated by relationship length; hence, Machiavellianism demonstrated a greater influence on the incidence of pretending orgasm behavior of women in long-term relationships.

Detection and Consequences

Previous research indicates that men place considerable importance on their partner's orgasm (McKibben, Bates, Shackelford, Hafen, & LaMunyon, 2010) and may therefore question partners about their experience or attend to likely orgasm cues (e.g., vocalizations, breathing rate). Despite men's interest in their partner's orgasm, detection rates (i.e., the ability to detect when a woman is pretending) appear to be low. Indeed only 55% of men report that they can recognize when their partner is pretending to experience orgasm (Mialon, 2012). Furthermore, men report that their partner pretends to experience orgasm less frequently than their partners report engaging in

this behavior (Ellsworth & Bailey, 2013). Though this form of deception is often successful, the consequences can be substantial. Indeed, the reactions of men who become aware that their partners have pretended to experience orgasm are similar to men responding to a partner's infidelity (e.g., anger, betrayal; Shackelford, Leblanc, & Drass, 2000). Hence, though pretending to experience orgasm is a form of deception frequently performed by women, there is an inherent risk that it may threaten the stability and integrity of the relationship. At present, there is a paucity of information assessing women's responses to a male partner pretending to experience orgasm or responses to this deception in casual relationships.

Infidelity

In some circumstances, the sexual act itself may constitute deception. Romantic relationships are typically formed with the expectation of romantic and sexual exclusivity. Infidelity is however widespread. It is most likely to occur when the primary partner does not meet important needs such as sex, intimacy, or companionship (Lewandowski & Ackerman, 2006), and may have a range of negative consequences including anger or violence, feelings of distress and betrayal, and termination of the relationship. Men and women are of course sensitive to this form of deception and though often conceptualized as a "dark" emotion, jealousy serves an important adaptive function. Jealousy is experienced in response to a real or imagined threat to a valued relationship and prompts men and women to respond to that threat. In sexual or romantic relationships, it may promote men and women to engage in mate retention behaviors which may strengthen the existing relationship, deter rivals, or encourage the partner to remain faithful. Factors influencing the type of mate retention behavior selected (e.g., enhancing the relationship or threatening a rival) include the quality or length of the relationship and attractiveness of the self, partner, or rival (e.g., Brewer & Riley, 2009).

Men and women do not differ with regard to the frequency or intensity of jealousy experienced and both men and women respond jealously to emotional (e.g., sharing personal feelings) and sexual (e.g., sexual intercourse with another person) infidelity. Men and women do differ with regard to the threats that elicit jealous behavior, and these differences reflect the specific evolutionary pressures experienced by each sex. Female infidelity places their male partners at risk of cuckoldry (i.e., unknowingly raising another man's child) resulting in a loss of important time and resources and potentially a loss in social status. Hence, men display greater distress in response to sexual infidelity than to emotional infidelity. Women, as the sex which experiences pregnancy and childbirth, do not face the risk of cuckoldry. The greater threat is posed by dissolution of the romantic relationship and diversion of resources to another woman. Therefore, women report greater distress in

response to emotional infidelity (e.g., Buss, Larsen, Westen, & Semelroth, 1992; Edlund & Sagarin, 2017).

It is important to note that research in this area has been criticized on both methodological and theoretical grounds. Most commonly, it is criticized for employing hypothetical forced choice scenarios in which participants are asked to imagine that their partner has been emotionally (but not sexually) unfaithful or sexually (but not emotionally) unfaithful and then report the scenario which is most upsetting. It is argued that women assume that men engaging in emotional infidelity would not do so without also being sexually unfaithful. Similarly, it is suggested that men assume women that are sexually unfaithful would not do so without emotional infidelity. Hence, critics of forced choice scenario research argue that men and women report sexual and emotional infidelity to be the most distressing respectively because they believe that the other form of infidelity must also be present. Despite these criticisms, sex differences have been identified using other methodological approaches. For example, Sagarin, Becker, Guadagno, Nicastle, and Millevoi (2003) report similar sex differences in response to sexual and emotional infidelity when using the aforementioned forced choice scenarios and when utilizing a continuous response scale (e.g., rating the likely distress on a scale of 1–7). Furthermore, sex differences have been identified in response to actual infidelity. For example, when faced with an unfaithful partner, men are more likely to interrogate their partner about sexual aspects of the relationship whereas women are more likely to ask questions about emotional elements of the relationship. When confronted about their own infidelity men are more likely to deny emotional aspects of the relationship while women are more likely to deny a sexual relationship (Kuhle, Smedley, & Schmitt, 2009).

Sexting

In other circumstances, it is not the sexual act which constitutes deception but the lack of a sexual act. Sexting, which may be defined as sending sexually explicit text messages, photographs, or videos to another person, is relatively common. For example, Delevi and Weisskirch (2013) report that 89% of college students had engaged in sexting. The use of deception within sexting is relatively widespread, with 48% of active sexters reporting having deceived their partner during sexting (Drouin et al., 2014). For example, partners may lie about what they are wearing or doing. Research suggests that deception during sexting is more common among women than men (i.e., similar to pretending to experience orgasm during sexual intercourse), with 45% of women and 24% of men reporting that they had lied while sexting their partner (Drouin et al., 2014). Consistent with the pretending orgasm research, individuals often deceive in order to improve the partner's experience though self-serving deception also occurred.

Conclusion

To conclude, deception is frequently employed by men and women when trying to obtain sex or during sexual activity. When deceiving to obtain sex, men and women appeal to the type of partner and relationship preferred by potential partners. For example, men may feign relationship commitment. Deception during sex is commonly characterized by pretending to experience orgasm, for example through the use of vocalizations, body movements, or an altered breathing rate. This is most frequent among women and is often performed in order to reassure the partner and strengthen the relationship. Though deception for and during sex is relatively common and performed by both men and women, it can lead to conflict, distress, and relationship dissolution.

References

Adamopoulou, E. (2013). New facts on infidelity. *Economics Letters, 121,* 458–462.

Alterovitz, S. S. R., & Mendelsohn, G. A. (2009). Partner preferences across the lifespan: Online dating by older adults. *Psychology and Aging, 24,* 513–517.

Bancroft, J., Loftus, J., & Long, J. S. (2003). Distress about sex: A national survey of women in heterosexual relationships. *Archives of Sexual Behavior, 32,* 193–208.

Brewer, G., Abell, L., & Lyons, M. (2016). Machiavellianism, pretending orgasm, and sexual intimacy. *Personality and Individual Differences, 96,* 155–158.

Brewer, G., & Hendrie, C. A. (2011). Evidence to suggest that copulatory vocalizations in women are not a reflexive consequence of orgasm. *Archives of Sexual Behavior, 40,* 559–564.

Brewer, G., & Riley, C. (2009). Height, relationship satisfaction, jealousy, and mate retention. *Evolutionary Psychology, 7,* 477–489.

Brym, R. J., & Lenton, R. L. (2001). *Love online: A report on digital dating in Canada.* Retrieved December 21, 2016, from http://modernloveonline.com/wp-content/uploads/loveonline.pdf.

Buss, D. M., Larsen, R. J., Westen, D., & Semmelroth, J. (1992). Sex differences in jealousy: Evolution, physiology, and psychology. *Psychological Science, 3,* 251–255.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100,* 204–232.

Clark, R. D., & Hatfield, E. (1989). Gender differences in receptivity to sexual offers. *Journal of Psychology and Human Sexuality, 2,* 39–55.

Cochran, S. D., & Mays, V. M. (1990). Sex, lies and HIV. *New England Journal of Medicine, 332,* 774–775.

Cooper, E. B., Conner, B. T., & Fauber, R. L. (2010). *Motivations for faking orgasm as predictors of sexual functioning and satisfaction.* Paper Session Presented at the LGBT Special Interest Group Sponsored Symposium at the Meeting of the Association for Behavioral and Cognitive Therapies, San Diego, CA.

Darling, C. A., & Davidson, J. (1986). Enhancing relationships: Understanding the feminine mystique of pretending orgasm. *Journal of Sex and Marital Therapy, 12,* 182–196.

Delevi, R., & Weisskirch, R. S. (2013). Personality factors as predictors of sexting. *Computers in Human Behavior, 29,* 2589–2594.

Desiderato, L. L., & Crawford, H. J. (1995). Risky sexual behavior in college students: Relationships between number of sexual partners, disclosure of previous risky behaviour, and alcohol use. *Journal of Youth and Adolescence, 24,* 55–68.

Dimoulas, E., Wender, S., Keenan, J. P., Gallup, G., & Goulet, N. (1998). Patterns of deception in human mating strategy. *Journal of Psychology and the Behavioural Sciences, 12,* 38–42.

Docan-Morgan, T., & Docan, C. A. (2007). Internet infidelity: Double standards and the differing views of women and men. *Communication Quarterly, 55,* 317–342.

Drouin, M., Tobin, E., & Wygant, K. (2014). "Love the way you lie": Sexting deception in romantic relationships. *Computers in Human Behavior, 35,* 542–547.

Edlund, J. E., & Sagarin, B. J. (2017). Chapter five—Sex differences in jealousy: A 23-year retrospective. *Advances in Experimental Social Psychology, 55,* 259–302.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication, 11,* 415–441.

Ellsworth, R. M., & Bailey, D. H. (2013). Human female orgasm as an evolved signal: A test of two hypotheses. *Archives of Sexual Behavior, 42,* 1545–1554.

Fales, M. R., Frederick, D. A., Garcia, J. R., Gildersleeve, K. A., Haselton, M. G., & Fisher, H. E. (2016). Mating markets and bargaining hands: Mate preferences for attractiveness and resources in two national U.S. studies. *Personality and Individual Differences, 88,* 78–87.

Fletcher, G. J. O., Kerr, P. S. G., Li, N. P., & Valentine, K. A. (2014). Predicting romantic interest and decisions in the very early stages of mate selection: Standards, accuracy, and sex differences. *Personality and Social Psychology Bulletin, 40,* 540–550.

Gibbs, J. L., Ellison, N. B., & Heino, R. D. (2006). Self-presentation in online personals: The role of anticipated future interaction, self-disclosure, and perceived success in Internet dating. *Communication Research, 33,* 152–177.

Hald, G. M., & Høgh-Olesen, H. (2010). Receptivity to sexual invitations from strangers of the opposite gender. *Evolution and Human Behavior, 31,* 453–458.

Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication, 59,* 367–386.

Haselton, M. G., Buss, D. M., Oubaid, V., & Angleitner, A. (2005). Sex, lies, and strategic interference: The psychology of deception between the sexes. *Personality and Social Psychology Bulletin, 31,* 3–23.

Hendrie, C. A., Brewer, G., Lewis, H., & Mills, F. (2014). Contemporary and historical evidence to suggest that women's preference for age at birth of first child remains consistent across time. *Archives of Sexual Behavior, 43,* 1373–1378.

Hitsch, G. J., Hortacsu, A., & Ariely, D. (2009). *What makes you click: An empirical analysis of online dating* (Unpublished manuscript). University of Chicago.

Jenkins, T. G., Aston, K. I., Pflueger, C., Cairns, B. R., & Carrell, D. T. (2014). Age-associated sperm DNA methylation alterations: Possible implications in offspring disease susceptibility. *PLoS Genetics, 10,* e1004458.

Johnson, A. K., Barnacz, A., Constantino, P., Triano, J., Shackelford, T. K., & Keenan, J. P. (2004). Female deception detection as a function of commitment and self-awareness. *Personality and Individual Differences, 37,* 1417–1424.

Kaighobadi, F., Shackelford, T. K., & Weekes-Shackelford, V. A. (2012). Do women pretend orgasm to retain a mate? *Archives of Sexual Behavior, 41,* 1121–1125.

Kamble, S., Shackelford, T. K., Pham, M., & Buss, D. M. (2014). Indian mate preferences: Continuity, sex differences, and cultural change across a quarter of a century. *Personality and Individual Differences, 70,* 150–155.

Keenan, J. P., Gallup, G., & Falk, D. (2003). *The face in the mirror.* New York: HarperCollins.

Keenan, J. P., Gallup, G. G., Goulet, N., & Kulkarni, M. (1997). Attributions of deception in human mating strategies. *Journal of Social Behavior and Personality, 12,* 45–52.

Knox, D., Schacht, C., Holt, J., & Turner, J. (1993). Sexual lies among university students. *College Student Journal, 27,* 269–272.

Kuhle, B. X., Smedley, K. D., & Schmitt, D. P. (2009). Sex differences in the motivation and mitigation of jealousy-induced interrogations. *Personality and Individual Differences, 46,* 499–502.

Lewandowski, G. W., & Ackerman, R. A. (2006). Something's missing: Need fulfilment and self-expansion as predictors of susceptibility to infidelity. *The Journal of Social Psychology, 146,* 389–403.

Lloyd, E. A. (2005). *The case of the female orgasm: Bias in the science of evolution.* Cambridge, MA: Harvard University Press.

Lo, S. K., Hsieh, A. Y., & Chiu, Y. P. (2013). Contradictory deceptive behavior in online dating. *Computers in Human Behavior, 29,* 1755–1762.

Lucchetti, A. E. (1999). Deception in disclosing one's sexual history: Safe-sex avoidance or ignorance? *Communication Quarterly, 47,* 300–314.

McCoy, M. G., Welling, L. M., & Shackelford, T. K. (2014). Development and initial psychometric assessment of the reasons for pretending orgasm inventory. *Evolutionary Psychology, 13,* 129–139.

McKibben, W. F., Bates, V. M., Shackelford, T. K., Hafen, C. A., & LaMunyon, C. W. (2010). Risk of sperm competition moderates the relationship between men's satisfaction with their partner and men's interest in their partner's copulatory orgasm. *Personality and Individual Differences, 49,* 961–966.

Mialon, H. M. (2012). The economics of faking ecstasy. *Economic Inquiry, 50,* 277–285.

Muehlenhard, C. L., & Shippee, S. K. (2010). Men's and women's reports of pretending orgasm. *Journal of Sex Research, 46,* 552–567.

Pawlowski, B., & Dunbar, R. I. M. (1999). Withholding age as a putative deception in mate search tactics. *Evolution and Human Behavior, 20,* 53–69.

Peterson, A., Carmen, R., & Gehr, G. (2013). Ovulatory shifts in mating intelligence. *Journal of Social, Evolutionary, and Cultural Psychology, 7,* 66–75.

Pew Research Center. (2016). *15% of American adults have used online dating sites or mobile dating apps.* Retrieved from http://www.pewinternet. org/2016/02/11/15-percent-of-american-adults-have-used-online-dating-sites-or-mobile-dating-apps/.

Rochadiat, A. M., Tong, S. T., & Novak, J. M. (2017). Online dating and courtship among Muslim American women: Negotiating technology, religious identity, and culture. *New Media & Society, 20*(4), 1618–1639. https://doi.org/10.1177/1461444817702396.

Sagarin, B. J., Becker, D. V., Guadagno, R. E., Nicastle, L. D., & Millevoi, A. (2003). Sex differences (and similarities) in jealousy: The moderating influence of infidelity experience and sexual orientation of the infidelity. *Evolution and Human Behavior, 24,* 17–23.

Schmitt, D. P., & International Sexuality Description Project. (2003). Universal sex differences in the desire for sexual variety: Tests from 52 nations, 6 continents, and 13 islands. *Journal of Personality and Social Psychology, 85,* 85–104.

Sedgewick, J. R., Flath, M. E., & Elias, L. J. (2017). Presenting your best self(ie): The influence of gender on vertical orientation of selfies on Tinder. *Frontiers in Psychology, 8,* 604.

Shackelford, T. K., Leblanc, G. J., & Drass, E. (2000). Emotional reactions to infidelity. *Cognition and Emotion, 14,* 643–659.

Singh, D. (2002). Female mate value at a glance: Relationship of waist-to-hip ratio to health, fecundity and attractiveness. *Neuroendocrinology Letters, 23,* 81–91.

Stebleton, M. J., & Rothenberger, J. H. (1993). Truth or consequences: Dishonesty in dating and HIV/AIDS-related issues in a college-age population. *Journal of American College Health, 42,* 51–54.

Steiger, S., Eichinger, T., & Honeder, B. (2009). Can mate choice strategies explain sex differences? The deceived persons' feelings in reaction to revealed online deception of sex, age, and appearance. *Social Psychology, 40,* 16–25.

Steiner, A. E. (1981). *Pretending orgasm by men and women: An aspect of communication in relationships* (Unpublished doctoral dissertation). California School of Professional Psychology, Berkeley, CA.

Thornhill, R., Gangestad, S. W., & Comer, R. (1995). Human female orgasm and mate fluctuating asymmetry. *Animal Behaviour, 50,* 1601–1615.

Toma, C. L., & Hancock, J. T. (2010). Looks and lies: The role of physical attractiveness in online dating self-presentation and deception. *Communication Research, 37,* 335–351.

Toma, C. L., & Hancock, J. T. (2012). What lies beneath: The linguistic traces of deception in online dating profiles. *Journal of Communication, 62,* 78–97.

Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34,* 1023–1036.

Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man* (pp. 136–179). Chicago: Aldine.

Ward, J. (2017). What are you doing on Tinder? Impression management on a matchmaking mobile app. *Information, Communication & Society, 20,* 1644–1659.

Waynforth, D., & Dunbar, R. I. M. (1995). Conditional mate choice strategies in humans: Evidence from lonely hearts' advertisements. *Behaviour, 132,* 755–779.

Whitty, M. T. (2002). Liar, liar! An examination of how open, supportive, and honest people are in chat rooms. *Computers in Human Behavior, 18,* 343–352.

Wiederman, M. W. (1997). Pretending orgasm during sexual intercourse: Correlates in a sample of young adult women. *Journal of Sex and Marital Therapy, 23,* 131–139.

Williams, S. S. (2001). Sexual lying among college students in close and casual relationships. *Journal of Applied Social Psychology, 31,* 2322–2338.

Zilbergeld, B. (1999). *The new male sexuality*. New York: Bantam.

Managing Face in the Midst of Interpersonal Deception: A Cross-Cultural Examination

Tara Suwinyattichaiporn and Mark A. Generous

Deception is a common communication act that people engage in daily. DePaulo, Kashy, Kirkendol, Wyer, and Epstein (1996) found that people engage in deceptive communication, on average, once or twice per day. Researchers have examined several aspects of deceptive communication, including types of deception and frequency (DePaulo, Ansfield, & Bell, 1996), the medium of deception (Whitty & Carville, 2008), whether most deceptive communications are spontaneous or planned (Whitty, Buchanan, Joinson, & Meredith, 2012), deception detection rate (Levine, Shaw, & Shulman, 2010), and the development of interpersonal deception theory (Buller & Burgoon, 1996). However, there is a lack of research that focuses on the communicative responses that occur immediately when the deception is detected. Therefore, this study adopts the communication concept of facework behaviors (Oetzel & Ting-Toomey, 2003) in order to understand the differing communication that happens in response to deception detection. In other words, the key question of this research is: How do people engage in facework behaviors once they have detected deception?

Interestingly, scholars have found that people from different cultures perceive and respond to deception differently (Bond & Atoum, 2000; Seiter, Bruschke, & Bai, 2002). For instance, individuals who belong to collectivistic

T. Suwinyattichaiporn (✉)
California State University, Fullerton, CA, USA
e-mail: tsuwinyattichaiporn@Fullerton.edu

M. A. Generous
California State Polytechnic University, Pomona, CA, USA
e-mail: magenerous@cpp.edu

567

cultures are likely to be more accepting of deceptive acts compared to those from individualistic cultures (Seiter et al., 2002). In order to observe cultural differences and their subsequent impact on responses to deception detection, the present study examines Thai individuals' and American individuals' self-construal identifications, as well as the ways in which expressed self-construal influences the type of facework behaviors that are enacted during a deception episode. The findings in this study contribute to the deceptive communication and interpersonal communication research in meaningful ways. First, this study expands our understanding of facework behaviors to the context of deception. Facework behaviors have been studied extensively (Oetzel, Ting-Toomey, Yokochi, Masumoto, & Takai, 2000; Oetzel et al., 2001; Ting-Toomey, 1985, 1988; Ting-Toomey et al., 1991), however, scholars have not fully examined how individuals engage in facework behaviors during a deceptive episode, particularly, in response to deception detection. It is important to examine this context because deception is universal and deceptive episodes are intuitively face-threatening and unpleasant situations. Detecting deception calls for individuals, regardless of the culture they are from, to engage in face management, thus establishing a warrant for the current study. Additionally, this study contributes to broadening the theoretical scope of interpersonal deception theory (IDT; Buller & Burgoon, 1996) by studying the responses to deception detection cross-culturally. Currently, there is a dearth of research that examines interpersonal deception cross-culturally, and this study provides a better understanding of how culture, particularly self-construal orientation, plays a role in the facework responses to deception. Moreover, it validates IDT's key assumptions that communicators are goal-oriented and receivers are active participants in a deceptive episode, which can help strengthen IDT's explanatory power.

LITERATURE REVIEW

Documented Differences in Thai and American Cultures

Cross-cultural research is paramount in cultivating an understanding of the process of enculturation as it pertains to communication phenomenon such as deception (Castillo & Mallard, 2012; Kim, Kam, Sharkey, & Singelis, 2008). As previously mentioned, the current investigation seeks to understand cultural influences on facework enactment following deception detection between Thai and American individuals. The decision to focus on Thai and American cultures was made for three primary reasons. First, one of the primary differences between American and Thai cultures is nested in the concept of individualism—cultures that value the success and prosperity of the individual—and collectivism—cultures that strive to benefit the family or social groups to which people collectively belong (Triandis, 1995). Indeed, previous research has documented the individualistic tendencies of American culture (Rokeach, 1979), as well as the emphasis on social harmony and

politeness present in Thai culture (Komin, 1990). Subsumed within this documented bifurcation are a myriad of behavioral distinctions between these two cultures, such as the way that intergenerational interactions occur, the amount of family-related communication present in the home, communication in the educational context, and the frequency of all communication patterns in general (Knutson, Hwang, & Vivatananukul, 1995; Rhee, Chang, & Rhee, 2003). Overall, the literature suggests that the collectivistic nature of Thai culture results in lower levels of communication across contexts, when compared to their American counterparts.

In addition, the concept of high- versus low-context cultures is documented in much of the cross-cultural literature in communication studies (Chua & Gudykunst, 1987; Kim, Pan, & Park, 1998; Würtz, 2005). Briefly defined, high-context cultures are those in which the denotative meaning of messages is less important than the environment, relationship, and social setting in which it occurs. In a low-context culture, verbal directness and linguistic cues are the most important form of meaning (Kim et al., 1998). Notably, high- and low-context cultures coincide with collectivistic and individualistic cultures, respectively (Gudykunst et al., 1996). Therefore, high-context communication is prominent in Thai culture, whereas low-context communication is more common in the American culture.

Adding to the difference between two cultures is the documented level of communication apprehension. Communication apprehension is defined as the level of anxiety associated with communicating with others (McCroskey, 1977). Perhaps linked to the collectivistic tendencies of Thai culture, it should come as little surprise that their desire to talk is much less than that of Americans (Knutson, Komolsevin, Chatiketu, & Smith, 2002; Verluyten, 1997). The reason for a significantly lower level of communication has been speculated to stem from increased levels of in-group sensitivity (Kim, Aune, Kim, Hunter, & Kim, 2001), inherent introversion (Sallinen-Kuperinen, 1986), and need to observe and acquire information before deciding to speak (Wiseman, 2002). Regardless of the interpretations of these behaviors, these cultural discrepancies appear to be at the root of the communicative patterns of these two cultures. Thus, as we begin to discuss the ways that these two cultures differ in their facework in the midst of deception, these significant distinctions are at the heart of the current investigation.

Understanding Deception

Deception has been studied in multiple contexts, ranging from close relationships (Jang, Smith, & Levine, 2002), police interrogations (Porter & Yuille, 1996), the ability to detect acts of deception (Levine, Park, & McCornack, 1999), and deceptive communication via computer-mediated technology (Wise & Rodriguez, 2013). These studies all rest under the assumption that deception is an intentional attempt to either conceal truthful information or

present false information with the goal of misleading a person or group of people (Rodgers, 2008). From the IDT standpoint, multiple goals are present in the mind of not only the deceiver but of the deception detector as well, including self, other, or relationship goals (Buller & Burgoon, 1996). Further, relational motivations, as well as instrumental and identity goals, can lead an individual to perform a deceptive act. This branch of assumptions requires the recognition that both parties involved in a deceptive episode possess agency and are capable of committing both intentional and unintentional acts of deception (Burgoon & Buller, 1994).

Indeed, the vast majority of work that is available on intercultural deception focuses either on "intercultural" as it pertains to the study of one non-Western culture (e.g., Yeung, Levine, & Nishiyama, 1999; Seiter, Weissman, Madrid, & Gass, 1990) or "cross-cultural" in reference to the individualistic/collectivistic differences between samples (e.g., Aune & Waters, 1994; Seiter et al., 2002). Although these differences are important to define and establish, more specific explications for the variances in cultural behaviors are present between American and Thai cultures. Accordingly, the present study intends to focus primarily on how self-construal and facework behaviors operate cross-culturally during a deceptive episode.

Self-construal and Responses to Deception

Self-construal is defined as "how individuals see the self in relation to others" (Cross, Hardin, & Gercek-Swing, 2011, p. 143). There are two major categories of self-construal: independent and interdependent self-construal. When someone has an independent self-construal, they tend to view themselves as more autonomous and self-serving. They care more about individual characteristics and uniqueness than group memberships. On the other hand, one who has an interdependent self-construal conforms to group desires and norms prior to considering their own. They care more about their relationships with others than their individual needs (Markus & Kitayama, 1991). Both dimensions of self-construal (i.e., independent and interdependent) exist on a continuum; that is, individuals are either low or high in independent and/or interdependent self-construal (Kim & Leung, 1997). Individuals who report having a high interdependent self-construal tend to be socialized in collectivistic cultures, while individuals who report having a high independent self-construal tend to be socialized in individualistic cultures (Gudykunst et al., 1996).

Self-construal is related to various communication behaviors. For instance, Yum (2004) found that despite collectivist associations with high accommodating behavior, self-construal orientation (as opposed to culture) was more accurate at depicting the ways that individuals enact behaviors of both accommodation and non-accommodation. Another study examined the importance of self-construal in times of stress or conflict. They found that individuals with a high interdependent self-construal will attempt to change the "self"

before attempting to change the situation that they find themselves in, whereas those with a high independent self-construal will attempt to manipulate their surroundings before modifying their own behavior (Weisz, Rothebaum, & Blackburn, 1984; Yang, 1986). In another study, Cross (1995) found that self-construal was the highest predictor of stress for international students, as well as for the behaviors that were used to cope with those stressful encounters. The aforementioned findings highlight the significant influence of individuals' self-construal on their behaviors in times of stress or discomfort.

When discussing self-construal and deception, it is important to understand how one's self-construal (independent and interdependent) influences perceptions of acts of deception (Levine et al., 2002). From an IDT perspective, the motivation behind a deceptive act (instrumental, relational, identity) is significantly associated with the frequency and intensity with which deceptive acts are enacted (Buller & Burgoon, 1996). Findings indicate that people from collectivistic cultures are likely to be more accepting of deceptive acts than people from individualistic cultures (Seiter et al., 2002). Given the strong correlations between collectivism and interdependent self-construal, as well as individualism and independent self-construal, it stands to reason that individuals' self-construal will be associated with reactions to deception detection. One key implication from these studies is the suggestion that where an individual's cultural system may value one set of behaviors, the self-construal of the people who belong to those cultures may dictate a different set of behaviors across a number of various contexts.

Facework and Deception

The concept of face stems from the idea that people, as social beings, desire to be acknowledged as worthy and desirable (Goffman, 1967). When confronted with a negative social predicament, many individuals opt to preserve their social status, their current relationship, or their own identity through a multitude of mitigating techniques (Goffman, 1967). Furthermore, individuals desire to not be socially attacked or interrogated, as these acts threaten and eventually cause us to "lose" our face (Ting-Toomey, 1985, 1988). It is the constant loss and saving of face that inspired the development of face negotiation theory (Ting-Toomey, 1988). This theory posits that individuals' cultural backgrounds (i.e., individualism vs. collectivism) influence the degree of self-oriented and other-oriented face-saving techniques they are more likely to perform (Ting-Toomey, 1988; Ting-Toomey et al., 1991). For example, Kim and Wilson (1994) found that people from individualistic cultures value the autonomous face (i.e., one's own image/face), while people from collectivistic cultures tend to enact other-face approval strategies (i.e., the other person's image in an interaction).

One situation in which facework is highly relevant is during deceptive episodes. There are several facework behaviors that become relevant

when discussing responses to deceptive communication. The three overar-ching categories of facework behaviors are (1) dominating behaviors, such as defending one's own opinion, expressing ones feeling, or using direct/passive aggression; (2) integrating behaviors, such as problem-solving or showing respect for opposing opinions; and (3) avoiding behaviors, such as apologizing, giving into the opposition, or pretending that the conflict does not exist (Ting-Toomey et al., 1991). These facework acts appear to be related to the target whose face the communicator wishes to save. For example, dominating behaviors are linked to self-face preservation, avoid-ing behaviors have been linked to other-face preservation, and integrating behaviors have been linked to mutual-face preservation (Oetzel, Garcian, & Ting-Toomey, 2008).

Previous research has revealed that deceptive acts are highly threatening to one's own face (Hodgins, Liebeskind, & Schwartz, 1996). Similar to a decep-tion detection episode, conflict interactions can also be quite face-threatening (Ting-Toomey et al., 1991). Cross-culturally, Cupach and Imahori (1993) found that Americans are more likely to enact direct and aggressive face-saving techniques, while Japanese are more likely to employ passive apolo-getic face-preserving acts in the episode of conflict. Accordingly, it has been posited that individualists are more likely to enact self-preserving responses to face-threatening acts (such as making excuses or justifications), whereas collectivists are more likely to enact self-deprecating techniques (such as acknowledging their lack of skill or effort) (Ting-Toomey & Kurogi, 1998). As it pertains to the actual behaviors within conflict episodes, individualists tend to enact more direct and potentially face-threatening behaviors (e.g., using the dominating or competing styles), whereas collectivists strive for more mutually face-saving acts (e.g., avoiding or obliging styles; Cocroft & Ting-Toomey, 1994).

There are two important points to take note of when discussing facework behaviors that may emerge during deception. First, it is acknowledged that deceivers engage in behavior that is both strategic (i.e., information is pur-posefully managed) and non-strategic (i.e., unintentional behavioral leakages occur) (Buller & Burgoon, 1996). Thus, when engaged in facework follow-ing deception, it is plausible that individuals will enact a variety of face-man-aging behaviors, both strategic and non-strategic. Second, as interpersonal deception is argued to be a process, it stands to reason that the cultural differ-ences between individuals (i.e., differences in self-construal) will contribute to their facework behaviors. Taking relevant literature into account, the follow-ing hypotheses are tested in the current study:

H1: Thais have higher interdependent self-construal than Americans whereas Americans have higher independent self-construal than Thais.

H2: Americans prefer dominating facework behaviors and Thais prefer avoid-ing facework behaviors in responses to deception detection.

H3: There is a *positive* association between interdependent self-construal and avoiding facework behaviors in a deception detection episode.

H4: There is a *negative* association between independent self-construal and avoiding facework behaviors in a deception detection episode.

Method

The present study used cross-sectional, self-report, survey research design to examine whether American and Thai participants score differently on self-construal orientations and whether they respond to a deceptive episode using different facework behaviors. Data from two countries were collected separately from June to August in 2014.

Participants

The sample consisted of 322 undergraduate students from a large Southwestern university in the US and a large university in Bangkok, Thailand. Of the 322 respondents, 143 were Thais and 181 were Americans. The average age for Americans was 23.63 ($M = 23.63$, SD $= 4.78$) and for Thais was 22.25 ($M = 22.25$, SD $= 3.76$). Of 143 Thai participants, 65 identified as female and 78 identified as male. Of 181 American participants, 97 identified as female, 82 identified as male, and 2 preferred not to disclose.

Procedure

This study employed convenience sampling procedures. Undergraduate students were recruited from communication courses at a university in Thailand and the US by their instructors to take the online survey in exchange for extra credit (under instructors' discretions). Students who wished to participate clicked a link that redirected them to the online survey (hosted on Qualtrics.com). First, they were presented with a participant information letter informing them about the study and that participation is voluntary. They were also informed that by clicking next, they have given consent for their participation. The survey consisted of three parts. The first part of the survey comprised demographic questions, including biological sex, age, and whether they are Thai or American. The second part included questions regarding their cultural values. In the final part, they read a vignette that represented a deceptive episode (see below) and responded to questionnaire items imagining what they would do in the scenario if they were Terry (the victim of deceptive act).

> Pat and Terry have been friends for 2 years. They take some classes together and sometimes hang out during the weekends. One Saturday night Terry invited Pat to come hang out with his new friends from the gym. Pat did not want to go

but does not know how to reject the invitation since Terry insists that he should come. As a result, Pat lies to Terry and says he promised to spend time with his parents who live in another province. Later that night, Terry accidentally sees Pat at the movie theatre so he sees that Pat is in town and not with his family in another province.

Measures

Self-Construal
Independent and interdependent self-construal orientations were measured by a 29-item Self-Construal Scale (Kim & Leung, 1997). This scale consisted of 29 questionnaire items: 15 items represent independent self-construal and 14 items represent interdependent self-construal. The participants responded on a 5-point Likert scale ranging from 1=strongly disagree to 5=strongly agree. Sample items include: "I voice my own opinion in group discussions," "I take responsibility for my own actions," and "I conceal my negative emotions so I won't cause unhappiness in my group." Both scale dimensions demonstrated strong internal consistency in the current investigation: independent self-construal scale (Cronbach's $\alpha = .83$) and interdependent self-construal (Cronbach's $\alpha = .83$).

Manipulation Check
Since the vignette method is adopted, manipulation check must be used to confirm participants' perceptions of the vignette. In this case, participants rated the seriousness of the deception episode by a five-item 5-point semantic differentials scale ranging from 1=not serious to 5=serious. Sample items include: "How serious was that lie?," "How evil was that lie?," and "How mean was that lie?"

Facework Behaviors
Questionnaire items for facework behaviors were derived from Oetzel, Garcia, and Ting-Toomey's facework study (2008). This particular scale represents three major types of facework behaviors including dominating, avoiding, and integrating behaviors. Participants responded on a 1 to 5 Likert-type scale from 1=extremely unlikely to 5=extremely likely. Sample items include "I directly expressed my feelings," "I tried to ask for help from another friend," and "I tried to defend my position." The scales demonstrate good validity and strong reliabilities in this study, with alpha coefficient ranging from .81 to .92.

RESULTS

Hypothesis 1 stated that Thais would score higher in interdependent self-construal than Americans, whereas Americans would score higher in independent self-construal than Thais. This hypothesis was supported. There

Table 30.1 Mean difference in independent and interdependent self-construal between Thais and Americans

Thai/American	Independent self-construal		Interdependent self-construal	
	M	SD	M	SD
Thais	3.17	.43	3.26	.44
Americans	4.04	.51	3.06	.61

Note Significant at the *p* < .001 level

was a significant difference in the scores of interdependent self-construal between Thais (*M*=3.26, SD=.44) and Americans (*M*=3.06, SD=.61); *t*(310)=3.40, *p*=.001, η^2=.04. In addition, there was a significant difference in the scores of independent self-construal between Thais (*M*=3.17, SD=.43) and Americans (*M*=4.04, SD=.51); *t*(316)=-6.56, *p*<.001, η^2=.12 (Table 30.1).

Hypothesis 2 stated that Americans prefer dominating facework behaviors and Thais prefer avoiding facework behaviors in responses to deception detection. Three independent-samples *t*-tests were conducted to investigate Thais and Americans' use of dominating, integrating, and avoiding facework behaviors. The hypothesis was supported. There was a statistically significant difference in the use of dominating and avoiding behaviors between the two groups. First, Americans scored higher in dominating behavior (*M*=3.08, SD=.76) compared to Thais (*M*=2.49, SD=1.05); *t*(322)=−5.95, *p*<.001, η^2=.10. Second, Americans scored lower in avoiding behavior (*M*=2.76, SD=.64) compared to Thais (*M*=3.30, SD=1.08); *t*(322)=5.57, *p*<.001, η^2=.09. The t-test on integrating behavior obtained a nonsignificant result, *t*(322)=−.446, *p*=.12. These results suggest that Thais and Americans do differ when it comes to responding to deception detection with regard to dominating and avoiding behaviors (Table 30.2).

Hypothesis 3 stated that there is a positive relationship between interdependent self-construal and avoiding facework behavior. This hypothesis was supported. A Pearson's product-moment correlation was used to analyze the data to find out the relationship between interdependent self-construal

Table 30.2 Mean difference in dominating and avoiding facework behaviors between Thais and Americans

Thai/American	Dominating		Avoiding	
	M	SD	M	SD
Thais	2.49	1.05	3.30	1.08
Americans	3.08	.76	2.76	.64

Note Significant at the *p* < .001 level

Table 30.3   Correlations between self-construal and three facework behaviors

Self-construal orientation	Dominating	Avoiding	Integrating
Independent self-construal	.09	−.15**	.01
Interdependent self-construal	−.02	.18**	.17**

Note Significance (1-tailed) **$p < .01$

and avoiding facework behavior. The findings indicate that interdependent self-construal was positively correlated with avoiding behaviors, $r(321) = .18$, $p < .01$. This suggests that people who have higher interdependent self-construal are more likely to engage in avoiding facework behavior when they detect deception.

Hypothesis 4 stated that there is a negative relationship between independent self-construal and avoiding facework behavior. This hypothesis was supported. A Pearson's product-moment correlation was utilized to analyze the relationship between independent self-construal and avoiding facework behaviors. The correlation coefficient $r(321) = -.15$, $p < .01$. indicates that scores of the two variables are significantly negatively correlated. The result suggests that the more independent self-construal people possess, the less likely they are to use avoiding facework behaviors in response to deception detection (Table 30.3).

DISCUSSION AND IMPLICATIONS

The purpose of the current study was to examine the following: (1) Thais and Americans self-construal orientations, (2) the differences between facework behaviors used by Thai and American participants upon detecting deception, and (3) the relationship between self-construal and three types of facework behaviors. The major findings of this study are:

1. Thais and Americans have different self-construal orientations. Thais have higher interdependent self-construal whereas Americans have higher independent self-construal.
2. When Thais detected deception (the victim in a deceptive episode), they tend to engage in avoiding facework behavior in response to the deceptive act.
3. When Americans detected deception, they tend to engage in dominating facework behavior in response to the deceptive act.
4. There is a positive correlation between independent self-construal and dominating facework behavior.
5. There is a positive correlation between interdependent self-construal and avoiding facework behavior.

Several findings lend support to the proposed hypotheses. To begin, results indicate that Thai participants report significantly higher interdependent self-construal compared to American participants; conversely, American participants report significantly higher independent self-construal compared to Thais. This is consistent with existing literature on both cultures (Komin, 1990; Triandis, 1993). Komin (1990) found that Thais were more collectivistic and Americans were more individualistic in their values and communication. Collectivism is closely related with interdependent self-construal, and individualism with independent self-construal. Gudykunst (2003) stated that individualists value personal rights, expressiveness, privacy, individual responsibility, personal goals, and voicing opinions more than collectivists. Meanwhile, collectivists value group harmony, collaboration, and maintaining face more than individualists. These descriptions demonstrate conceptual similarity between self-construal and individualism/collectivism. Therefore, it is not a surprising finding that Thai participants scored higher in interdependent self-construal and Americans scored higher in independent self-construal in this study.

In terms of Thais and Americans' use of facework behaviors in response to deception detection, results were also consistent with existing literature. Thai participants reported using avoiding behaviors the most when deception was detected. One of the questionnaire items for avoiding behaviors stated, "I tried to ask for help from another friend," which truly captures the non-confrontational nature of this type of communication. As previously documented, Thais are highly avoidant and polite in their communication (Knutson et al., 1995; Knutson et al., 2002). Thai culture is often referred to as high context (Adair & Brett, 2005; Chua & Gudykunst, 1987). Thus, it makes sense that Thais would prefer using avoiding facework behaviors, even when they learn that the other person is being deceptive.

Conversely, American participants reported using significantly more dominating facework behaviors compared to Thais. These findings indicate that Americans are significantly more likely to directly confront a friend when they detect that person has deceived them compared to Thais. This is consistent with the individualism and independent self-construal literature (Gudykunst, 2003; Triandis, 1993). Dominating facework behaviors consist of survey items such as "I let the other person know clearly what I was thinking" (Oetzel et al., 2008, p. 390). This scale item exemplifies someone directly voicing their opinion, which is a behavioral characteristic of individuals from individualistic cultures who report high independent self-construal.

This study also found a positive association between interdependent self-construal and avoiding facework behaviors. This means that people with higher interdependent self-construal tendencies (Thais—as found in this study) tend to engage in avoiding facework behaviors when they detect someone lying. This is consistent with past literature because people with a high

interdependent self-construal care more about group harmony over their personal goals (Markus & Kitayama, 1991). Avoiding behaviors are non-confrontational, thus, avoiding could better serve the mutual-face needs of all people involved and maintain group harmony.

On the other hand, a negative association was found between independent self-construal and avoiding behaviors, which indicates that a higher independent self-construal tends to be associated with a decreased likelihood to employ avoiding behaviors in response to deception detection. This finding supports existing literature on independent self-construal and facework behaviors (Cupach & Imahori, 1993; Markus & Kitayama, 1991). For instance, Cupach and Imahori (1993) found that, in general, Americans tend to use more dominating facework behaviors as compared to Japanese people. Consistently, Merkin (2018) links individualism to direct facework and collectivism to indirect facework. Since people with high independent self-construals are described as more autonomous and place high value in voicing their opinions (Markus & Kitayama, 1991), it is reasonable that they would not be avoiding the deceiver, especially when they have detected the deceptive act.

Limitations, Future Research, and Conclusion

The present study is not without limitations. First, there were more American participants than Thai participants. Second, the present study only measured one cultural orientation variable, which was self-construal. Future studies should consider larger sample sizes, different cultural backgrounds, and measure more than one cultural variable (e.g., high-context/low-context, power distance). In addition, future studies should examine cross-cultural elements of interpersonal deception theory, such as strategic and non-strategic behaviors. The present study focused on the differences in the use of facework behavior in response to a deceptive communication act, which highlighted only strategic behavior. Also, future researchers should consider using retrospective prompts by asking participants to think of a recent time they engaged in a deceptive act or detected deception. Since the present study utilized a vignette method in framing the deceptive episode, the results relied on participants being able to imagine themselves as the characters in the vignette. Finally, longitudinal diary studies could also help deception researchers understand cultural influences on how individuals respond to deception in relationships.

To conclude, this study examined whether culture plays a role in people's responses to a deceptive communication act. Specifically, it was found that Americans have high independent self-construal and prefer using dominating facework behaviors when they have detected deception. Thais have high interdependent self-construal and prefer using avoiding facework behaviors in response to deception detection. Current findings indicate that people's cultural backgrounds influence the ways in which they communicatively respond to deception detection.

REFERENCES

Adair, W. L., & Brett, J. M. (2005). The negotiation dance: Time, culture, and behavioral sequences in negotiation. *Organization Science, 16*(1), 33–51. https://doi.org/10.1287/orsc.1040.0102.

Aune, R. K., & Waters, L. L. (1994). Cultural differences in deception: Motivations to deceive in Samoans and North Americans. *International Journal of Intercultural Relations, 18*(2), 159–172. https://doi.org/10.1016/0147-1767(94)90026-4.

Bond, C. F., Jr., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin, 26*(3), 385–395.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18*(2), 155–184. https://doi.org/10.1007/BF02170076.

Castillo, P. A., & Mallard, D. (2012). Preventing cross-cultural bias in deception judgments: The role of expectancies about nonverbal behavior. *Journal of Cross-Cultural Psychology, 43*, 967–978. https://doi.org/10.1177/0022022111415672.

Chua, E. G., & Gudykunst, W. B. (1987). Conflict resolution styles in low-and high-context cultures. *Communication Research Reports*, *4*(1), 32–37.

Cocroft, B., & Ting-Toomey, S. (1994). Facework in Japan and the United States. *International Journal of Intercultural Relationships, 18*, 469–506. https://doi.org/10.1016/0147-1767(94)90018-3.

Cross, S. (1995). Self-construals, coping, and stress in cross cultural adaptation. *Journal of Cross-Cultural Psychology, 26*(6), 673–697. https://doi.org/10.1177/002202219502600610.

Cross, S. E., Hardin, E. E., & Gercek-Swing, B. (2011). The what, how, why, and where of self-construal. *Personality and Social Psychology Review, 15*(2), 142–179.

Cupach, W. R., & Imahori, T. T. (1993). Managing social predicaments created by others: A comparison of Japanese and American facework. *Western Journal of Communication, 57*(4), 431–444. https://doi.org/10.1080/10570319309374466.

DePaulo, B. M., Ansfield, M. E., & Bell, K. L. (1996). Theories about deception and paradigms for studying it: A critical appraisal of Buller and Burgoon's interpersonal deception theory and research. *Communication Theory, 6*(3), 297–310.

Goffman, E. (1967). On face-work. *Interaction Ritual*, 5–45.

Gudykunst, W. B. (2003). *Cross-cultural and intercultural communication*. Thousand Oaks, CA: Sage.

Gudykunst, W. B., Matsumoto, Y., Ting-Toomey, S., Nishida, T., Kim, K., & Heyman, S. (1996). The influence of cultural individualism-collectivism, self construals, and individual values on communication styles across cultures. *Human Communication Research, 22*(4), 510–543. https://doi.org/10.1111/j.1468-2958.1996.tb00377.x.

Hodgins, H. S., Liebeskind, E., & Schwartz, W. (1996). Getting out of hot water: Facework in social predicaments. *Journal of Personality and Social Psychology, 71*(2), 300–314. https://doi.org/10.1037/0022-3514.71.2.300.

Jang, S. A., Smith, S., & Levine, T. (2002). To stay or to leave? The role of attachment styles in communication patterns and potential termination of romantic relationships following discovery of deception. *Communication Monographs, 69*(3), 236–252. https://doi.org/10.1080/03637750216543.

Kim, M., Aune, K. S., Kim, H., Hunter, J. E., & Kim, J. (2001). The effect of culture and self-construal on predispositions toward verbal communication. *Human Communication Research, 27,* 382–408. https://doi.org/10.1111/j.1468-2958.2001.tb00786.x.

Kim, M. S., & Leung, T. (1997). *A revised self-construal scale* (Unpublished manuscript). University of Hawaii at Manoa.

Kim, D., Pan, Y., & Park, H. S. (1998). High-versus low-context culture: A comparison of Chinese, Korean, and American cultures. *Psychology and Marketing, 15*(6), 507–521. https://doi.org/10.1002/(SICI)1520-6793(199809)15:6<507:AID-MAR2>3.0.CO;2-A.

Kim, M. S., Kam, K. Y., Sharkey, W. F., & Singelis, T. M. (2008). "Deception: Moral transgression or social necessity?" Cultural-relativity of deception motivations and perceptions of deceptive communication. *Journal of International and Intercultural Communication, 1,* 23–50. https://doi.org/10.1080/17513050701621228.

Kim, M. S., & Wilson, S. R. (1994). A cross-cultural comparison of implicit theories of requesting. *Communication Monographs, 61*(3), 210–235. https://doi.org/10.1080/03637759409376334.

Knutson, T. J., Hwang, J. C., & Vivatananukul, M. (1995). A comparison of communication apprehension between Thai and USA student samples: Identification of different cultural norms governing interpersonal communication behaviors. *Journal of National Research Council of Thailand, 27*(1), 21–46.

Knutson, T. J., Komolsevin, R. V., Chatiketu, P., & Smith, V. R. (2002). A comparison of Thai and US American willingness to communicate. *Journal of Intercultural Communication Research, 31,* 3–12.

Komin, S. (1990). *Psychology of the Thai people: Values and behavioral patterns.* Research Center, National Institute of Development Administration.

Levine, T. R., Lapinski, M. K., Banas, J., Wong, N., Hu, A., Endo, K., … Anders, L. N. (2002). Self-construal, self and other benefit, and the generation of deceptive messages. *Journal of Intercultural Communication Research, 31*(1), 29–47.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs, 66*(2), 125–144. https://doi.org/10.1080/03637759909376468.

Levine, T. R., Shaw, A., & Shulman, H. C. (2010). Increasing deception detection accuracy with strategic questioning. *Human Communication Research, 36*(2), 216–231. https://doi.org/10.1111/j.1468-2958.2010.01374.x.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253. https://doi.org/10.1037/0033-295X.98.2.224.

McCroskey, J. C. (1977). Oral communication apprehension: A summary of recent theory and research. *Human Communication Research, 4,* 78–96. https://doi.org/10.1111/j.1468-2958.1977.tb00599.x.

Merkin, R. S. (2018). Individualism-collectivism applied to direct versus indirect facework. In *Saving Face in Business* (pp. 119–136). New York: Palgrave Macmillan.

Oetzel, J., Garcia, A. J., & Ting-Toomey, S. (2008). An analysis of the relationships among face concerns and facework behaviors in perceived conflict situations: A four-culture investigation. *International Journal of Conflict Management, 19*(4), 382–403. https://doi.org/10.1108/10444060810909310.

Oetzel, J. G., & Ting-Toomey, S. (2003). Face concerns in interpersonal conflict a cross-cultural empirical test of the face negotiation theory. *Communication Research, 30*(6), 599–624. https://doi.org/10.1177/0093650203257841.

Oetzel, J., Ting-Toomey, S., Masumoto, T., Yokochi, Y., Pan, X., Takai, J., & Wilcox, R. (2001). Face and facework in conflict: A cross-cultural comparison of China, Germany, Japan, and the United States. *Communication Monographs, 68*(3), 235–258. https://doi.org/10.1080/03637750128061.

Oetzel, J. G., Ting-Toomey, S., Yokochi, Y., Masumoto, T., & Takai, J. (2000). A typology of facework behaviors in conflicts with best friends and relative strangers. *Communication Quarterly, 48*(4), 397–419. https://doi.org/10.1080/01463370009385606.

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior, 20*(4), 443–458. https://doi.org/10.1007/BF01498980.

Rhee, S., Chang, J., & Rhee, J. (2003). Acculturation, communication patterns, and self-esteem among Asian and Caucasian American adolescents. *Adolescence, 38*(152), 749–768.

Rogers, R. (2008). *Clinical assessment of malingering and deception*. New York, NY: Guilford Press; Thousand Oaks, CA: Sage.

Rokeach, M. (1979). Change and stability in American value systems. In M. Rokeach (Ed.), *Understanding human values: Individual and societal* (pp. 1968–1971). New York: Free Press.

Sallinen-Kuparinen, A. (1986). Finnish communication reticence: Perceptions and self-reported behavior. *Studia Philologica Jyvaskylaensia, 19*. Jyvaskyla: University of Jyvaskyla.

Seiter, J. S., Bruschke, J., & Bai, C. (2002). The acceptability of deception as a function of perceivers' culture, deceiver's intention, and deceiver-deceived relationship. *Western Journal of Communication, 66*(2), 158–180. https://doi.org/10.1080/10570310209374731.

Seiter, J. S., Wiseman, R. L., Madrid, V., Gass, R. H., & Riggio, R. D. (1990, June). *Untangling webs: The effects of ethnicity on the detection of deception*. Paper Presented to the Intercultural Communication Division at the Annual Meeting of the International Communication Association, Dublin, Ireland.

Ting-Toomey, S. (1985). Toward a theory of conflict and culture. In W. B. Gudykunst, L. P. Stewart, & S. Ting-Toomey (Eds.), *Communication, culture, and organizational processes* (pp. 71–86). Beverly Hills, CA: Sage.

Ting-Toomey, S. (1988). Intercultural conflict styles: A face negotiation theory. In Y. Y. Kim & W. Gudykunst (Eds.), *Theories in intercultural communication* (pp. 47–92). Newberry Park, CA: Sage.

Ting-Toomey, S., Gao, G., Trubisky, P., Yang, Z., Soo Kim, H., Lin, S. L., & Nishida, T. (1991). Culture, face maintenance, and styles of handling interpersonal conflict: A study in five cultures. *International Journal of Conflict Management, 2*(4), 275–296. https://doi.org/10.1108/eb022702.

Ting-Toomey, S., & Kurogi, A. (1998). Facework competence in intercultural conflict: An updated face negotiation theory. *International Journal of Intercultural Relationships, 22*(2), 187–225. https://doi.org/10.1016/S0147-1767(98)00004-2.

Triandis, H. C. (1993). Collectivism and individualism as cultural syndromes. *Cross-Cultural Research, 27*(3–4), 155–180. https://doi.org/10.1177/106939719302700301.

Triandis, H. C. (1995). *Individualism & collectivism.* Boulder, CO: Westview Press.

Verluyten, S. P. (1997). *Some cultural aspects of Thai companies with recommendations for westerners.* Antwerp, Belgium: Center for ASEAN Studies.

Weisz, J. R., Routhbaum, F. M., & Blackburn, T. C. (1984). Standing out and standing in: The psychology of control in American and Japan. *American Psychologist, 39*, 955–969. https://doi.org/10.1037/0003-066x.39.9.955.

Whitty, M. T., Buchanan, T., Joinson, A. N., & Meredith, A. (2012). Not all lies are spontaneous: An examination of deception across different modes of communication. *Journal of the American Society for Information Science and Technology, 63*(1), 208–216.

Whitty, M. T., & Carville, S. E. (2008). Would I lie to you? Self-serving lies and other-oriented lies told across different media. *Computers in Human Behavior, 24*(3), 1021–1031.

Wise, M., & Rodriguez, D. (2013). Detecting deceptive communication through computer-mediated technology: Applying interpersonal deception theory to texting behavior. *Communication Research Reports, 30*(4), 342–346.

Wiseman, R. L. (2002). Intercultural communication competence. In W. B. Gudykunst & B. Moody (Eds.), *Handbook of international and intercultural communication* (pp. 207–224). Thousand Oaks, CA: Sage.

Würtz, E. (2005). Intercultural communication on web sites: A cross-cultural analysis of websites from high-context cultures and low-context cultures. *Journal of Computer-Mediated Communication, 11*(1), 274–299. https://doi.org/10.1111/j.1083-6101.2006.tb00313.x.

Yang, K. S. (1986). Chinese personality and its change. In M. H. Bong (Ed.), *Psychology of Chinese people* (pp. 107–170). Hong Kong: Oxford University Press.

Yeung, L. N., Levine, T. R., & Nishiyama, K. (1999). Information manipulation theory and perceptions of deception in Hong Kong. *Communication Reports, 12*(1), 1–11.

Yum, Y. O. (2004). Culture and self-construal as predictors of responses to accommodative dilemmas in dating relationships. *Journal of Social and Personal Relationships, 21*(6), 817–835. https://doi.org/10.1177/0265407504047839.

Lying Online: Examining the Production, Detection, and Popular Beliefs Surrounding Interpersonal Deception in Technologically-Mediated Environments

Catalina L. Toma, James Alex Bonus and Lyn M. Van Swol

Defined as the deliberate attempt of one person to generate a false belief in another, deception has always been a staple of human communication, with people lying frequently to achieve personal, social, economic, or political goals (Vrij, 2008). In recent decades, deception has taken on a new life with the advent and tremendous popularity of new communication technologies, such as email, texting, social network sites (SNSs), and online dating. New types of deception have emerged and entered the popular vernacular. For example, *phishing* refers to malicious websites enticing users to reveal their personal information by masquerading as trustworthy websites (e.g., banks) (see Toma, 2014). *Email spam* refers to unsolicited, commercial messages

C. L. Toma (✉) · L. M. Van Swol
Department of Communication Arts,
University of Wisconsin-Madison, Madison, WI, USA
e-mail: ctoma@wisc.edu

L. M. Van Swol
e-mail: vanswol@wisc.edu

J. A. Bonus
School of Communication, Ohio State University, Columbus, OH, USA
e-mail: bonus.1@osu.edu

583

that sometimes host malware or phishing links. *Astroturfing* is the practice of using fake online accounts and identities to give the impression that an organization, individual, or idea is more popular than it actually is (e.g., by leaving favorable comments or bestowing "likes"). In the realm of interpersonal communication, people's options for lying to relationship partners have broadened, as communication technologies have become widely used on an everyday basis. When planning their lies, people can now choose from an extensive array of media options for delivering deceptive messages. Similarly, they must contend with the possibility of encountering deception in media formats that differ substantially from traditional face-to-face (FtF) settings. This has created a maelstrom of concern over the deceptiveness of online communication, with people worried that technology is encouraging deception and eroding interpersonal trust (see Baym, 2010, for a review).

This chapter provides a critical review of the literature on technology and interpersonal deception, or the lies people tell using mediated communication within their personal relationships and social encounters (rather than large-scale, commercially or politically motivated online deceptions). First, we address the actual extent of *deception production* in technologically-mediated spaces: How much do people lie online and why? Then, we address *popular beliefs about deception*: How much deception do people think occurs in technologically-mediated spaces? Are these beliefs accurate or systematically biased? Finally, we discuss *deception detection*: Is it possible to detect mediated deception, using both human judges and computerized techniques? What cues are useful in these detection efforts?

For all three topics, we pay particular attention to the theoretical and empirical link between deception and technological features and affordances. Technological features refer to design elements such as the presence of photographs or the provision of real names that should impact deception. Technological affordances refer to abilities that users have for social interaction in technological spaces (see Treem & Leonardi, 2013). When it comes to deception, relevant affordances are (1) the reduction or elimination of nonverbal cues, since people and researchers alike associate deception with the production of diagnostic cues, such as an aversion of eye gaze or fidgeting (Global Deception Research Team, 2006); (2) unlimited composition time, or the ability to take as much time as desired for constructing deceptive messages; (3) synchronicity, or the extent to which social interaction takes place in real time, because many lies are elicited through question asking by the interaction partner; (4) editability, or the ability to revise messages until they come across as convincing; (5) distribution, or the ability to interact with others without sharing the same physical space, which should enable liars to dissemble about their current whereabouts and activities; and (6) recordability, or the extent to which interaction partners can preserve a copy of the messages exchanged and hence retain evidence of deception.

Online Deception Production

As a scholarly topic, deception production refers to the prevalence and types of lies people generate and to their motivations for doing so. When it comes to technologically-mediated communication, researchers have tended to examine deception production in one medium at a time, with only a handful of studies conducting multimedia comparisons. Below, we review studies on the individual media that have received the most scholarly attention (i.e., texting, SNSs, and online dating) and end with the studies that have conducted comparisons across the media.

Deception in Texting

Texting, or the exchange of brief messages via mobile phones and corresponding desktop applications (e.g., iMessage, WhatsApp), has become one of the most frequently used media for social interaction, especially among teenagers and young adults (Lenhart, Purcell, Smith, & Zickuhr, 2010). Perhaps the most salient feature of texting, as indicated by its very name, is that it enables text-based interaction, with communication partners typing out their thoughts without having to manage their own or scrutinize others' non-verbal behaviors.

Several studies have investigated the prevalence of deception in college students' text messages (Birnholtz, Guillory, Hancock, & Bazarova, 2010; Reynolds et al., 2011; Reynolds, Smith, Birnholtz, & Hancock, 2013; Smith, Hancock, Reynolds, & Birnholtz, 2014). These studies' methodology involved a two-pronged approach. First, participants were asked to retrieve 15–50 of their most recent text messages, identify any lies in these messages, and rate the magnitude of these lies, from slightly deceptive to extremely deceptive. Second, coders were trained to categorize lies. Results are remarkably consistent across studies: About 10% of texts were reported to be deceptive, with the magnitude of deception slightly below the midpoint of the scale (i.e., between 2.5 and 3 on a scale from 1 [slightly deceptive] to 5 [extremely deceptive]). A fifth to a third of texting lies were coded as pertaining to the initiation, conclusion, or coordination of social interactions (e.g., "on my way," when one has not yet left the house; "sorry can't talk now, gotta work" when one doesn't in fact have to work; "can't wait to see you" when one is not in fact excited about an upcoming meeting). These lies were aptly labeled *butler lies*, because they serve the function of managing one's availability for current or future interpersonal contact with one's interlocutor, in the same way that a butler would. Qualitative analyses revealed that participants told butler lies in order to extricate themselves from unwanted interactions without offending or hurting conversation partners. These unwanted interactions are especially likely to take place via texting, because texting

generates expectations that users will be constantly connected and perpetually available. Thus, butler lies represent a way of coping with the social demands for availability imposed by texting, and are told in order to maintain relationships.

The final study in this series (Smith et al., 2014) examined the *distribution* of texting lies across participants in order to ascertain whether deception was a common behavior for all text users (i.e., the *everyday liars perspective*, see DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996) or whether it was driven by a select group of users who lied a great deal (i.e., the *prolific liars perspective*, see Serota, Levine, & Boster, 2010). Support for both perspectives emerged. On the one hand, 77% of all the participants sent at least one deceptive text, supporting the *everyday liars perspective*. On the other hand, the distribution of texting lies had a pronounced positive skew, with 5% of the sample having told 15% of all the texting lies, supporting the *prolific liars perspective*. Notably, almost half of all the texts sent by these prolific liars were deceptive.

As described, the literature has specifically tied the prevalence of butler lies with the affordances of texting that generate expectations for constant connectivity (i.e., ease of use and wide adoption). Additionally, researchers have postulated that texting lies are enabled by the location, activity, and temporal ambiguities inherent to the texting medium (Smith et al., 2014). In other words, users can easily generate lies about where they are, what they are doing, and when they have received a message because this information is unavailable to communication partners.

Deception in SNSs

The defining characteristic of SNSs (e.g., Facebook, Instagram, LinkedIn) is that they connect users with typically large networks of friends, family members, acquaintances, employers, and even strangers. Thus, one of the key affordances of SNSs is publicness—but a special type of publicness, where at least some of the audience members know the users well and can verify their claims. This type of publicness can be theoretically expected to keep deception in check. Additionally, SNSs offer affordances that should facilitate the production of deception, such as unlimited composition time (allowing users to come up with suitable lies) and the opportunity to edit messages (allowing users to craft statements that appear credible).

Given this set of affordances, how deceptive are SNS profiles? Back et al. (2010) theorized two competing possibilities. According to the extended real-life hypothesis, SNS profiles should accurately portray users' personalities, because users do not want to come across as deceptive in front of audiences who know them well. According to the idealized virtual-identity hypothesis, users should take advantage of unlimited composition time and editability in order to compose embellished versions of self that can impress audiences.

The data supported the first contention: Unacquainted observers were able to correctly ascertain users' personalities based on a simple perusal of SNS profiles, suggesting that these profiles accurately displayed profile owners' personalities. Publicness, then, appears to curb deception in SNS profiles. A follow-up study addressed this question from the profile owners' perspective: How accurately do SNS users think they come across in their profiles (Toma & Carlson, 2015)? Some results were consistent with Back and colleagues' findings in that SNS users believed their profiles represented them accurately on a plethora of dimensions of self (e.g., "physically attractive," "creative," "likeable"), but better than reality on others (e.g., "calm," "relaxed," "adventurous," "funny"). This pattern provides evidence of users' strategic approach to SNS deception: Publicness indeed curbed deception on issues that were easily verifiable by the audience, such as users' physical attractiveness, but not on more subjective issues, such as how "cool" or humorous users were. For the latter, users appear to have strategically taken advantage of SNS affordances that increase control over self-presentational claims (i.e., editability and asynchronicity) in order to create flattering, yet believable images of self.

The studies above focused on Facebook or similar SNSs, where users connect with friends and acquaintances for social purposes. SNSs dedicated to professional networking are another arena where the issue of deception is highly salient. One study has investigated the deceptiveness of online resumes on LinkedIn compared to that of traditional, pen-and-paper resumes (Guillory & Hancock, 2012). Results show that both online and paper resumes contained about three lies on average, supporting the notion that people lie routinely in everyday life to accomplish personal goals. While there was no difference in overall deception between the two resume formats, online resumes contained fewer lies about work experiences and responsibilities (i.e., matters that could be easily verified by their audience, which typically contains past employers and colleagues), but more lies about personal interests (i.e., subjective matters that are difficult to verify). As was the case earlier with Facebook, the affordances of online resume websites also appear to encourage strategic deception. Specifically, publicness curbs deception about issues that can be verified by the audience, whereas editability and unlimited composition time encourage deception about more subjective issues, allowing users to present more flattering versions of self than warranted by reality after all.

Deception in Online Dating

In recent years, online dating has become one of the premier venues for finding romantic partners, for both short- and long-term relationships. Online dating sites operate by asking users to describe themselves via personal profiles and then enabling them to browse through the site's database of other

singles in order to identify romantic interests. The first point of contact between interested parties happens through messaging on the site; afterward, daters can connect through other media and FtF at their convenience (Toma, 2015). Online dating deception can take place at the profile self-presentation stage and also during direct interactions with potential mates. To our knowledge, the literature has focused exclusively on profile deception. Studies are needed to examine the use of deceptions during initial interactions between online daters.

The process of profile composition is facilitated by editability and unlimited composition time, two affordances that allow online daters a great degree of control over crafting their statements—certainly much greater than what is available FtF, where daters must think on their feet and cannot take back gaffes or awkward behavior. Dating, whether online or offline, produces high evaluative concerns, as daters voluntarily subject themselves to deep scrutiny from potential mates and open themselves to the possibility of rejection. Therefore, having control over one's claims, as is the case online, should be psychologically reassuring and should be used strategically to convey images of self that impress potential mates and are less likely to be rejected. Deception is a plausible strategy for achieving a desirable self-presentation; therefore, it should be facilitated by editability and unlimited composition time.

However, other affordances might temper the use of deception. Notably, the anticipation of future interaction, a defining characteristic of online dating, renders lies discoverable as the budding relationship progresses. Lies about physical appearance can be observed as soon as the first date, and lies about education, employment, or relational history are likely to be exposed down the line. Deception is regarded by many as a deal-breaker in the establishment of intimate relationships (Ellison, Hancock, & Toma, 2012), which is why online daters interested in developing such relationships should use it with caution. Furthermore, recordability allows online daters to retain copies of partners' profiles and thus preserve evidence of deception. Faced with this evidence, online daters may not justify their deceptions as misunderstandings or memory lapses. In sum, the anticipation of future interaction and recordability can be expected to act as powerful deception constraints.

The push and pull between deception enablers (asynchronicity, editability) and constraints (anticipation of future interaction, recordability) in heterosexuals' online dating profiles was examined in a comprehensive project by Toma, Hancock, and Ellison. A pattern of frequent but subtle deception emerged, consistent with theoretical predictions that online daters should take advantage of deception enablers to construct a more desirable self, while being mindful of deception constraints in order to not alienate potential partners. About 80% of online daters lied either about their height, weight, or age, measured objectively by the authors in the lab. However, the lies were small and strategic. On average, men added about an inch to their height to cater to women's stereotypical preferences for taller men; women subtracted about

eight pounds from their weight to cater to men's stereotypical preferences for thinner women; and age was relatively honestly portrayed across the board, arguably because age is not a malleable characteristic or one about which one can claim ignorance (unlike height and weight), and therefore age deceptions cannot be as readily justified (Toma, Hancock, & Ellison, 2008). Photographs were the most deceptive profile element, perhaps in response to the subjectivity of the medium: As a static and two-dimensional representation of self, a photograph cannot render an embodied person in the same way as FtF interaction does. This subjectivity provided online daters with more leeway for misrepresentation by selecting especially attractive photographs, engaging in carefully orchestrated posing, or using software to rectify flaws. Gender differences in photographic deception also emerged, with women engaging in more embellishment of their physical appearance, presumably to cater to the higher premium men put on women's attractiveness. When compared to photographs taken by the researchers in the lab, women's photographs were rated as more deceptive by a group of judges, and more likely to contain discrepancies related to age, skin quality, and hairstyles. Women's photographs were taken, on average, 17 months prior to composing the profile, whereas men's were taken, on average, only six months prior. Thus, women's photographs depicted a younger and potentially more attractive version of self (Hancock & Toma, 2009). Finally, the strategic aspect of online dating deception was highlighted by the fact that daters who were rated as less physically attractive were more likely to enhance their attractiveness by posting deceptive photographs and lying about physical descriptors (height, weight, age) than their more attractive counterparts; however, daters' physical attractiveness did not affect their use of deception on profile elements unrelated to physical appearance (income and occupation). Thus, online daters used deception in a limited way to rectify shortcomings, rather than lying indiscriminately simply because technology makes lying effortless (Toma & Hancock, 2010).

This pattern of deception suggests that online daters balance out deception enablers and constraints by lying judiciously and in a highly calculated manner. Several additional studies support this contention. Online daters reported lying less when seeking serious, long-term relationships than casual, short-term ones, arguably because deception is more likely to be detected and more injurious to relational prospects in the former than latter case (Gibbs, Heino, & Ellison, 2006). Men reported lying more about their physical appearance and personality characteristics when anticipating to meet a potential partner via email, where lies are less detectable, than via FtF, where lies are more detectable (Guadagno, Okdie, & Kruse, 2012). Finally, a large survey of over 5000 heterosexual online daters provided additional evidence of the gendered nature of online deception, with men reporting having lied more about characteristics valued by women (e.g., personal assets and relationship goals), and women reporting having lied more about characteristics valued by men (e.g., weight) (Hall, Park, Song, & Cody, 2010).

Deception Across the Media

Several studies have taken on the task of comparing rates of deception across the media in order to shed light on how technological affordances, which vary across the media, might result in different patterns of deception production. The seminal paper on this topic investigated college students' rates of self-reported deception via the phone, email, instant messenger (IM) and FtF communication, using the feature-based model as a theoretical guide (Hancock, Thom-Santelli, & Ritchie, 2004). According to this framework, the media contain the following features, a term used synonymously with affordances, that facilitate deception: (1) distribution, which allows users to dissemble about their whereabouts and activities; (2) synchronicity, which has been shown to increase deception because most lies emerge spontaneously in conversation; and (3) recordlessness, which ensures that the liars' statements cannot be preserved and used as evidence of their dishonesty. According to this framework, the phone contains the most features that facilitate deception (distribution, synchronicity, and recordlessness), whereas email contains the fewest (only distribution); IM (containing distribution, near-synchronicity, and near-recordlessness) and FtF (containing recordlessness and synchronicity) are tied for second place, with a fairly equal number of features that facilitate deception. This exact pattern was observed in the data: Most lies were told via the phone, followed by IM and FtF, and least via email, providing support to the feature-based approach. Notably, participants in this study reported engaging in deception in approximately a quarter of their social interactions, on average, consistent with the *everyday liars perspective*.

However, these findings did not replicate several years later. In two subsequent studies, George and Robb (2008) found no significant differences in deception among these four channels of communication, although they did find the overall rate of participants' daily deception (about a quarter of all social interactions) to be the same. The latter findings indicate that factors unrelated to affordances, such as habitual use of the media, may also play a part in deception production.

The studies above (Hancock et al., 2004; George & Robb, 2008) focused on dyadic media, or media used for one-on-one interaction. Another line of research has extended this scope to include broadcast media, or media that allow for one-to-many interaction, such as chat rooms and SNSs. To explain deception patterns across this broader range of media, Warkentin, Woodworth, Hancock, and Cormier (2010) introduced warranting theory, which proposes that a medium's deceptiveness should be inversely proportional with the number of warrants it supports. Warrants are defined as connections between online and offline claims, and are theorized to reduce deception because they render online claims more verifiable. For example, the inclusion of a personal name or of a photograph where the user is visually identifiable should make a participant in a chat room more accountable for her claims.

Similarly, links with social networks, which the authors labeled acquaintance warrants, are predicted to increase honesty because people don't want to lie in front of audiences who can detect this deception (see also the discussion of deception in SNSs). In an examination of deceptive behaviors in email, IM, online forums, chat rooms, and SNSs, the authors found that the presence of more warrants in a medium was correlated to fewer lies and, notably, fewer *serious* lies. Acquaintance warrants were the most powerful in reducing deception, indicating the premium people put on coming across as trustworthy to their social networks.

One final study (Drouin, Miller, Wehle, & Hernandez, 2016) examined the prevalence of deception across four online venues: social media, online dating, anonymous chat rooms, and sexual websites. Although these authors did not specifically utilize warranting theory, their results were consistent with it: Participants reported being least deceptive on social media sites, which include acquaintance warrants, followed by online dating sites, which do not include acquaintance warrants, but are expected to generate these warrants when daters meet potential partners FtF. Participants reported being most deceptive in chat rooms and sexual websites, which are devoid of acquaintance warrants.

Summary and Conclusions

Several overarching insights emerge from the growing literature on the production of interpersonal deception in technologically-mediated settings. First, deception appears to be a commonplace phenomenon, as it is in FtF settings (see DePaulo et al., 1996). Lies are told on a routine basis and, similarly to FtF interaction, a few prolific liars may be responsible for a disproportionate number of these lies (see Serota et al., 2010). Second and relatedly, there is robust evidence that mediated deception is a strategic act. From online daters who embellish themselves subtly and for the purpose of rectifying specific shortcomings, to SNS users who restrict their deception so as to not come across as liars to their acquaintance networks, to texting users who concoct butler lies because these lies are both polite and virtually undetectable, people lie in technologically-mediated settings cautiously and with awareness of their relational goals. Contrary to popular beliefs, a point we will detail in the next section, people do not lie online simply because it is easy to do so. Third, the literature has made great strides in postulating relationships between technological features and affordances and the production of deception. In particular, warrants (e.g., photographs, network connections) have been theorized to be a technological feature that should strongly affect deception production, and so have numerous affordances (i.e., the reduction of nonverbal cues, editability, unlimited composition time, distribution, and recordability). However, more theorizing is needed to specify (a) whether all features and affordances are equally potent in affecting deception production; (b) how

these features and affordances work together, given that, in everyday practice, they are bundled together in unique configurations in each medium; and (c) and how features and affordances intersect with users' psychological characteristics (e.g., self-monitoring, generalized trust), relational goals and motivations, and other contextual factors to shape deception. So far, the theories advanced to explicate mediated deception (e.g., the feature-based model, warranting theory) have tended to be feature- and affordance-centric, without detailing how these technological aspects work together with factors unrelated to the technology. This broader theorizing is needed in the future.

Popular Beliefs About Online Deception

Early reasoning about technology use provided a rather dismal outlook on its implications for social relationships. Generally, these arguments suggested that mediated interactions would be less fulfilling and of lower quality than comparable FtF interactions (Walther & Parks, 2002), and that technology would have a pervasive, monolithic, and detrimental impact on interpersonal relationships (Baym, 2010). Although such predictions have been repeatedly refuted in the literature (for a review, see Walther & Parks, 2002), these negative biases permeate public folk theories about mediated communication (e.g., Toma, Jiang, & Hancock, 2016; Drouin et al., 2016), and appear to trigger greater expectations of deceit in certain mediated environments. For example, surveys have shown that most people suspect deceit to be prevalent online (e.g., Toma et al., 2016; Caspi & Gorsky, 2006), particularly in environments that allow for more anonymity and less visibility (e.g., online chat rooms, sexual websites) than in environments that are less anonymous and more visible (e.g., social media profiles; Drouin et al., 2016). However, even online dating websites—which people generally use in hopes of making offline social connections—are characterized by suspicion. Indeed, people expect others to lie about a range of topics, including their age, height, weight, employment status, smoking behavior, and more (Ellison et al., 2012). One national survey found that such deceit was ranked as the single biggest drawback of meeting partners online (Smith & Duggan, 2013). Importantly, across all online environments, people expect others to engage in more deceptive acts than they engage in themselves, independent of actual rates of deception (Toma et al., 2016). In fact, a large number of people report engaging in their own deceptive acts online *because* they believe everybody else is already doing so (Drouin et al., 2016).

One useful explanation for these disparities is the "self-other asymmetry," a well-documented social bias suggesting that people view themselves more favorably than others, and they judge others more harshly than themselves (Pronin, Gilovich, & Ross, 2004). Because deception is generally regarded as an undesirable behavior (Vrij, 2008), and because people have an innate drive to view themselves positively

(Pronin et al., 2004), they distance themselves from deception, but do not provide the same courtesy to others. In fact, the self-other asymmetry gap widens when people are asked about the prevalence of deception in mediated (rather than FtF) environments, which suggests that people have unique biases about the deceptive qualities inherent to online interactions (Toma et al., 2016).

Generally, most research regarding individuals' beliefs about deception has focused on online text-based environments with real (or imagined) audiences. Less research has examined beliefs about other types of media (e.g., text messaging and phone calls) that are often used within close interpersonal relationships. Mediated interactions within close relationships are characterized by varied reductions in nonverbal cues (e.g., text-only instant messages, audio-only phone calls, and audiovisual video chat interactions). Because prior research has shown that people tend to rely on visual and auditory nonverbal cues (e.g., gaze aversion and pitch) when formulating opinions about possible instances of deception in FtF settings (Global Deception Research Team, 2006), the absence of some (or most) of these cues simultaneously engenders shifts in peoples' beliefs about deception.

For example, studies that have compared interpersonal interactions in text-only, audio-only, or audiovisual mediated environments have generally found that greater access to nonverbal cues fosters greater perceptions of credibility (Burgoon, Blair, & Strom, 2008) and trust between people (Bos, Olson, Gergle, Olson, & Wright, 2002) than when fewer cues are available, even in circumstances where people are instructed to analyze messages for deceit. Researchers argue that because deceivers tend to focus their efforts on crafting honest nonverbal demeanors, they are best able to persuade others of their honesty in mediated environments that provide access to the greatest number of such cues (Burgoon et al., 2008).

Beyond perceptions about available cues, it is also possible that people harbor unique beliefs about the mere *decision* to communicate over the media rather than FtF. This issue remains underexplored in the current literature. Indeed, survey research suggests that media selection is often an active, rational choice, particularly in close relationships (Frisby & Westermann, 2010). In the context of romantic relationships, the perceived ambiguity of CMC is associated with heightened suspicion that one's romantic partner is purposefully distancing themselves from an encounter by capitalizing on the asynchronous affordances of text-based media. Consistent with this perception, some types of negative biases toward relational partners (i.e., "they're purposefully ignoring me") have been shown to increase during mediated encounters (Scissors, Roloff, & Gergle, 2014). Such findings echo the results of qualitative work suggesting that trust is impaired in cue-deficient CMC environments, which limits the communication of emotion and can lead partners to misinterpret each others' motives (Herlein & Ancheta, 2014).

However, two studies to date have contradicted the notion that people are more suspicious of CMC (rather than FtF) interactions (Bonus & Toma, in preparation), even when people are explicitly given an option about how they would prefer to interact (Van Swol & Braun, 2014). These studies have consistently reported no differences in the levels of suspicion or trust resulting from interactions that occur over CMC or FtF. Taken together, these findings suggest that the negative biases permeating public attitudes about online communication with anonymous audiences are not equivalently reflected in their attitudes about mediated communication within close personal relationships. Although more research is undoubtedly required, the sheer ubiquity of mediated interactions with relational partners in everyday life has likely modified the way in which these more personal interactions are perceived relative to the types of less personal interactions that characterize online environments.

ONLINE DECEPTION DETECTION

People may think it would be more difficult to detect deception online. After all, one has fewer social cues to judge the sender in this leaner environment because one may lack nonverbal and parasocial information. Further, much communication online is asynchronous, so senders have time to craft a more believable reply (Toma & Hancock, 2012), but a deceptive sender communicating synchronously FtF has the burden of monitoring their own behavior for apparent truthfulness (Hartwig, Granhag, Stromwall, 2007; Levine et al., 2011), monitoring their partner for suspicion (Burgoon, Buller, Dillman, & Walther, 1995; Duran, Hall, McCarthy, & McNamara, 2010), and then responding to any perceived suspicion in real time. Thus, one might expect people to have a fairly good rate of detecting deception FtF, but this has not been supported by research. Research on FtF communication has found that, on average, people are able to detect deception at levels just slightly above chance (Aamodt & Custer, 2006; Bond & DePaulo, 2006). In fact, most nonverbal cues that people use to judge whether a source is being deceptive have low predictive reliability (Vrij, 2008).

Other research suggests that detecting deception online might actually be more accurate than detecting deception FtF. First, less synchronous forms of communication allow the receiver to spend more time examining the message (Whitty, Buchanan, Joinson, & Meredith, 2012). Also, online communication is more likely to be written. Thus, the receiver is not focused on nonverbal cues, which may not be a useful tool to detect deception, and can focus on other more reliable cues like consistency of the message, the level of detail, and sensory information provided (Logue, Book, Frosina, Huizinga, & Amos, 2015; Park, Levine, McCornack, Morrison, & Ferrerra, 2002). Other research has found that people who perceive themselves as unskilled deceivers are more likely to deceive online than FtF compared to people who

perceive that they are more skilled at deception (Van Swol & Paik, 2017). Thus, people who might avoid deception FtF because of concerns about their ability to pull off the deception may be willing to try to lie online. The bar to producing lies may be lower online, and this may increase the number of low-skill or transparent liars using online communication to deceive. Similarity, Dunbar, Jensen, Tower, and Burgoon (2014) suggested that leaner channels of communication may actually hamper the ability of skilled deceivers to deceive undetected. Finally, people are more suspicious of online communication and may be more inclined to pay attention to cues that the message is deceptive rather than mindlessly accept the message as truthful (Caspi & Gorsky, 2006; Van Swol, Braun, & Kolb, 2015; Van Swol & Paik, 2017; Whitty & Carville, 2008).

Researchers have started to examine these competing accounts on the influence of technology on deception detection. Several studies have found that participants are better able to detect deception through online communication (i.e., email or text chat) than FtF communication (Foglia, 2015; Van Swol & Braun, 2014; Van Swol et al., 2015). Although they did not compare online to FtF communication, Burgoon et al. (2008) found that the detection of deceptive messages was more accurate for the leaner channels of written and audio-only than audiovisual communication. However, Van Swol and Paik (2017) found no differences in detection rates of online versus FtF deception. All of these studies used written communication for online communication. Foglia (2015) used pre-created email communication that the receiver could read but not respond with questions, but Van Swol and Braun (2014) and Van Swol et al. (2015) used more synchronous and interactive text chat communication in which the receiver could ask the sender questions in real time. This distinction between synchronous and asynchronous online communication is important because Burgoon, Chen, and Twitchell (2010) found that synchronous online communication fostered more interactivity and trust, but that deceptive senders could use that interactivity to their advantage. They noted that "The combination of deception and synchronous communication may, however, be the most dangerous because deceivers can create a pseudo-relationship when interactivity is high and can capitalize on the truth bias that is more pronounced under real-time conditions" (p. 363). Thus, within online communication there are several differences in affordances that affect the detection of deception.

Van Swol and Braun (2014) and Van Swol and Paik (2017) allowed the sender to decide which channel to use to communicate (FtF or text chat), but Foglia (2015) and Van Swol et al. (2015) randomly assigned communication to either online or FtF. This difference—random assignment or decision by sender—is worth exploring because the decision to use a leaner channel to communicate, especially when a richer channel is readily available, conveys important information to the receiver, and the receiver may try to determine what motivated the sender to communicate in a leaner channel.

In contrast to the above studies, Hancock, Woodworth, and Goorha (2010) found that when liars are highly motivated, receivers are less likely to accurately detect deception online than FtF. They suggested that nonverbal cues to deception may be more salient and accurate under high stakes situations (DePaulo & Kirkendol, 1989; DePaulo, Kirkendol, Tang, & O'Brien, 1988; DePaulo, Lanier, & Davis, 1983; Toma & Hancock, 2010), and that senders may be more motivated to spend time designing a believable message with higher stakes. The former would increase detection accuracy for FtF, and the later would decrease it for leaner communication channels like written online communication (Hancock et al., 2010). Research has found that highly motivated senders are less likely to be detected when only communicating verbally (DePaulo et al., 1983). In addition, other research with fairly high stakes lies (Dunbar et al., 2015) has found that professional interviewers were better able to detect deception FtF than through videoconferencing, but in this case both forms of communication were verbal and synchronous. Thus, while the research is sparse, the majority of studies, which were all low stakes situations, found that receivers were more likely to detect deception through online communication than FtF (Foglia, 2015; Van Swol & Braun, 2014; Van Swol et al., 2015) or found no differences (Van Swol & Paik, 2017). Yet these results should not be generalized to high stakes situations, where the opposite pattern emerged (Hancock et al., 2010). More research on low and high stakes deception is needed.

Given the increasing amount of communication occurring through online channels and research indicating that people are more likely to deceive online than FtF (e.g., Naquin, Kurtzberg, & Belkin, 2010; Van Swol & Braun, 2014; Zimbler & Feldman, 2011), researchers and practitioners have examined tools to ferret out deception online. To leverage the written nature of most online communication, research has focused on psycholinguistic cues (e.g., Duran et al., 2010; Newman, Pennebaker, Berry, & Richards, 2003). Psycholinguistics examines how aspects of language can reflect psychological behavior. For example, liars tend to use language that reflects more negative emotions, more distancing, and less sensory and perceptual detail (for review see, Hauch, Blandón-Gitlin, Masip, & Sporer, 2015). Researchers have suggested using these tools for online deception (e.g., Ho, Hancock, Booth, & Liu, 2016). However, many linguistic indicators of deception have small effect sizes (Hauch et al., 2015) and would not be useful for a one-time determination of deception. Therefore, for everyday judgments of interpersonal deception online, a better indicator may be determining whether the information presented is verifiable outside the online communication. Further, someone using deception online may have more reticence and avoid presenting information that could be easily contradicted either by evidence or later offline conversation (Toma & Hancock, 2010).

Conclusions

Mediated interaction spaces differ from traditional FtF communication by providing users with features and affordances that impact how much they lie, what they believe about others' lies, and the extent to which they are able to detect deception. Thus, these spaces allow scholars to formulate more precise specifications about the mechanics of deception. If people believe that lying can be betrayed by nonverbal cues, how much do they lie when these cues are not available for scrutiny, as is the case online? How much do they think others lie? If nonverbal cues are not in fact diagnostic of deception, will people's deception detection ability be improved when these cues are unavailable?

This chapter has summarized the sizeable body of research that has accumulated in the past several decades on interpersonal deception and technology. In a nutshell, this research reveals that technological spaces present users with features and affordances that both facilitate (i.e., deception enablers: unlimited composition time and editability) and inhibit deception (i.e., deception constraints: recordability, publicness, and the anticipation of future interaction). Online liars appear adaptable, taking advantage of deception enablers while being mindful of deception constraints, in order to construct strategic lies. Nonetheless, when it comes to judging others' online deceptions, people hold simplistic and biased beliefs—they appear to disregard deception constraints, imagining that online deception is much more rampant than it actually is. These biases appear to be rooted in a belief that others are less ethical and less strategic than oneself, lying online simply because it is technologically easy to do so. Deception detection in mediated spaces is perhaps the least understood. A handful of studies indicate that deception detection might be superior online than FtF, at least in the case of small lies, arguably because online communicators are not led astray by non-diagnostic nonverbal cues and because they have more attentional resources to scrutinize messages. But this advantage might not extend to consequential lies. Finally, a growing body of research is showing how deception can be detected by leveraging the text-based nature of online communication: A small number of linguistic cues are useful for classifying messages as truthful or deceptive, with relatively high accuracy. Without a doubt, communication technology offers a rich platform for examining and re-examining the mechanics of interpersonal deception.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15,* 6–11.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., et al. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science, 21,* 372–374.

Baym, N. (2010). *Personal connections in a digital age.* Cambridge, UK: Polity Press.

Birnholtz, J., Guillory, J., Hancock, J., & Bazarova, N. (2010, February). On my way: Deceptive texting and interpersonal awareness narratives. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (pp. 1–4). ACM.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234.

Bos, N., Olson, J., Gergle, D., Olson, G., & Wright, Z. (2002, April). Effects of four computer-mediated communications channels on trust development. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 135–140). ACM.

Bonus, J. A., & Toma, C. L. *Sex, lies, and media use: Perceptions of deception about romantic infidelity across face-to-face and mediated communication.* Manuscript in preparation.

Burgoon, J. K., Blair, J. P., & Strom, R. E. (2008). Cognitive biases and nonverbal cue availability in detecting deception. *Human Communication Research, 34,* 572–599.

Burgoon, J. K., Buller, D. B., Dilman, L., & Walther, J. B. (1995). Interpersonal deception. *Human Communication Research, 22,* 163–196.

Burgoon, J. K., Chen, F., & Twitchell, D. P. (2010). Deception and its detection under synchronous and asynchronous computer-mediated communication. *Group Decision and Negotiation, 19,* 345–366.

Caspi, A., & Gorsky, P. (2006). Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior, 9*(1), 54–59.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995.

DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 51–70). Dordrecht, The Netherlands: Kluwer.

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivational impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12,* 177–202.

DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45,* 1096–1103.

Drouin, M., Miller, D., Wehle, S. M., & Hernandez, E. (2016). Why do people lie online? "Because everyone lies on the internet". *Computers in Human Behavior, 64,* 134–142.

Dunbar, N. E., Jensen, M. L., Burgoon, J. K., Kelley, K. M., Harrison, K. J., Adame, B. J., et al. (2015). Effects of veracity, modality, and sanctioning on credibility assessment during mediated and unmediated interviews. *Communication Research, 42,* 649–667.

Dunbar, N. E., Jensen, M. L., Tower, D. C., & Burgoon, J. K. (2014). Synchronization of nonverbal behaviors in detecting mediated and non-mediated deception. *Journal of Nonverbal Behavior, 38,* 355–376.

Duran, N. D., Hall, C., McCarthy, P. M., & McNamara, D. S. (2010). The linguistic correlates of conversational deception: Comparing natural language processing technologies. *Applied Psycholinguistics, 31,* 439–462.

Ellison, N. B., Hancock, J. T., & Toma, C. L. (2012). Profile as promise: A framework for conceptualizing veracity in online dating self-presentations. *New Media & Society, 14,* 45–62.

Foglia, V. (2015). *Detecting deception in face to face and computer mediated conversations* (Unpublished thesis). Algoma University. Retrieved from http://archives.algomau.ca/main/node/42532.

Frisby, B. N., & Westerman, D. (2010). Rational actors: Channel selection and rational choices in romantic conflict episodes. *Journal of Social and Personal Relationships, 27,* 970–981.

George, J. F., & Robb, A. (2008). Deception and computer-mediated communication in daily life. *Communication Reports, 21*, 92–103.

Gibbs, J. L., Ellison, N. B., & Heino, R. D. (2006). Self-presentation in online personals: The role of anticipated future interaction, self-disclosure, and perceived success in internet dating. *Communication Research, 33,* 1–26.

Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37*, 60–74.

Guadagno, R. E., Okdie, B. M., & Kruse, S. A. (2012). Dating deception: Gender, online dating, and exaggerated self-presentation. *Computers in Human Behavior, 28,* 642–647.

Guillory, J., & Hancock, J. T. (2012). The effect of LinkedIn on deception in resumes. *Cyberpsychology, Behavior, and Social Networking, 15,* 135–140.

Hall, J. A., Park, N., Song, H., & Cody, M. J. (2010). Strategic misrepresentation in online dating: The effects of gender, self-monitoring, and personality traits. *Journal of Social and Personal Relationships, 27,* 117–135.

Hancock, J., Thom-Santelli, J., & Ritchie, T. (2004). Deception and design: The impact of communication technology on lying behavior. In E. Dykstra-Erickson & M. Tscheligi (Eds.), *Proceedings of the 2004 Conference on Human Factors in Computing Systems* (pp. 129–134). New York: Association for Computing Machinery.

Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication, 59,* 367–386.

Hancock, J. T., Woodworth, M. T., & Goorha, S. (2010). See no evil: The effect of communication medium and motivation on deception detection. *Group Decision and Negotiation, 19,* 327–343.

Hartwig, M., Granhag, P. A., & Stromwall, L. A. (2007). Guilty and innocent suspects' strategies during interrogations. *Psychology, Crime, & Law, 13,* 213–227.

Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review, 19,* 307–342.

Hertlein, K. M., & Ancheta, K. (2014). Clinical application of the advantages of technology in couple and family therapy. *The American Journal of Family Therapy, 42,* 313–324.

Ho, J. M., Hancock, J. T., Booth, C., & Liu, X. (2016). Computer-mediated deception: Strategies revealed by language-action cues in spontaneous communication. *Journal of Management Information Systems, 33,* 393–420.

Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (2010). Social media & mobile internet use among teens and young adults. *Pew Internet & American Life Project.* Washington, DC. Retrieved from https://files.eric.ed.gov/fulltext/ED525056.pdf.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., et al. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research, 37,* 377–403.

Logue, M., Book, A. S., Frosina, P., Huizinga, T., & Amos, S. (2015). Using reality monitoring to improve deception detection in the context of the cognitive interview for suspects. *Law and Human Behavior, 39,* 360–367.

Naquin, C. E., Kurtzberg, T. R., & Belkin, L. Y. (2010). The finer points of lying online: E-mail versus pen and paper. *Journal of Applied Psychology, 95,* 387–394.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665–675.

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrerra, M. (2002). How people really detect lies. *Communication Monographs, 69,* 144–157.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review, 111,* 781–799.

Reynolds, L., Gillette, S., Marder, J., Miles, Z., Vodenski, P., Weintraub, A., … Hancock, J. T. (2011). Contact stratification and deception: Blackberry messenger versus SMS use among students. In P. Hinds (Ed.), *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (pp. 221–224). New York, NY: ACM.

Reynolds, L., Smith, M. E., Birnholtz, J. P., & Hancock, J. T. (2013, February). Butler lies from both sides: Actions and perceptions of unavailability management in texting. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (pp. 769–778). ACM.

Scissors, L. E., Roloff, M. E., & Gergle, D. (2014, April). Room for interpretation: The role of self-esteem and CMC in romantic couple conflict. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3953–3962). ACM.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36,* 2–25.

Smith, A. W., & Duggan, M. (2013). *Online dating & relationship.* Washington, DC: Pew Research Center. Retrieved from http://pewinternet.org/Reports/2013/Online-Dating.aspx.

Smith, M. E., Hancock, J. T., Reynolds, L., & Birnholtz, J. (2014). Everyday deception or a few prolific liars? The prevalence of lies in text messaging. *Computers in Human Behavior, 41,* 220–227.

Toma, C. L. (2014). Political phishing. In K. Harvey (Ed.), *Encyclopedia of social media and politics.* Thousand Oaks, CA: Sage.

Toma, C. L. (2015). Online dating. In C. R. Berger & M. E. Roloff (Eds.), *International encyclopedia of interpersonal communication.* Chichester: Wiley-Blackwell.

Toma, C. L., & Carlson, C. L. (2015). How do Facebook users think they come across in their profiles?: A meta-perception approach to Facebook self-presentation. *Communication Research Reports, 32,* 93–101.

Toma, C. L., & Hancock, J. T. (2010). Looks and lies: The role of physical attractiveness in online dating self-presentation and deception. *Communication Research, 37,* 335–351.

Toma, C. L., & Hancock, J. T. (2012). What lies beneath: The linguistic traces of deception in online dating profiles. *Journal of Communication, 62,* 78–97.

Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34,* 1023–1036.

Toma, C. L., Jiang, L. C., & Hancock, J. T. (2016). Lies in the eye of the beholder: Asymmetric beliefs about one's own and others' deceptiveness in mediated and face-to-face communication. *Communication Research, 45*(8), 1167–1192.

Treem, J. W., & Leonardi, P. M. (2013). Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association. *Annals of the International Communication Association, 36,* 143–189.

Van Swol, L. M., & Braun, M. T. (2014). Channel choice, justification of deception, and detection. *Journal of Communication, 64,* 1139–1159.

Van Swol, L. M., Braun, M., & Kolb, M. (2015). Deception, detection, demeanor, and truth bias in face-to-face and computer mediated communication. *Communication Research, 42,* 116–1142.

Van Swol, L. M., & Paik, J. E. (2017). Deciding how to deceive: Differences in communication and detection between good and bad liars. *Communication Quarterly, 65,* 503–522.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities.* Chichester, UK: Wiley.

Walther, J. B., & Parks, M. R. (2002). Cues filtered out, cues filtered in: computer-mediated communication and relationships. In M. L. Knapp & J. A. Daly (Eds.), *Handbook of interpersonal communication* (3rd ed., pp. 529–563). Thousand Oaks, CA: Sage.

Warkentin, D., Woodworth, M., Hancock, J. T., & Cormier, N. (2010). Warrants and deception in computer-mediated communication. In K. Inkpen (Ed.), *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (pp. 9–12). New York, NY: ACM.

Whitty, M. T., Buchanan, T., Joinson, A. N., & Meredith, A. (2012). Not all lies are spontaneous: An examination of deception across different modes of communication. *Journal of the American Society for Information Science and Technology, 63,* 208–216.

Whitty, M. T., & Carville, S. E. (2008). Would I lie to you? Self-serving lies and other oriented lies told across different media. *Computers in Human Behavior, 24,* 1021–1031.

Zimbler, M., & Feldman, R. S. (2011). Liar, liar, hard drive on fire: How media context affects lying behavior. *Journal of Applied Social Psychology, 41,* 2492–2507.

Contexts of Deceptive Communication: Groups and Organizations

Deceptive Communication in Group Contexts

Jeremy R. Winget and R. Scott Tindale

Prior to the 1960s, American manufactures dominated the automotive market. However, by the end of the decade, American automotive manufacturers' concerns began to rise with the increase of imported cars. Denying defeat, Ford Motor Company tried to remain competitive by producing the Pinto. Eager to have their subcompact ready, Ford condensed their typical drafting timeline, which meant any design changes that were typically made before production would instead be made during it (Shaw & Barry, 2001).

Before production, Ford crash-tested various prototypes, partially to learn whether they met a safety standard proposed by the National Highway Traffic Safety Administration (NHTSA) that aimed to reduce fires in traffic collisions. The standard required all new automobiles be able to withstand a rear-end impact of 20 miles per hour (mph) without fuel loss by 1972 (Shaw & Barry, 2001). When Ford crash-tested their Pinto prototypes, all failed the 20 mph test. Later, Ford crash-tested the final version of the Pinto and found the same result: ruptured gas tanks and dangerous leaks.

Ford knew that the Pinto represented a serious fire hazard when struck from the rear and faced a decision: (1) keep the existing design, thereby meeting the production timetable but jeopardizing consumer safety; or (2) delay production of the Pinto by redesigning the gas tank to make it safer and concede another year of subcompact dominance to foreign manufacturers. Ford ultimately pushed ahead with the original design and continued to

J. R. Winget (✉) · R. Scott Tindale
Department of Psychology, Loyola University Chicago, Chicago, IL, USA
e-mail: jwinget@luc.edu

R. Scott Tindale
e-mail: Rtindal@luc.edu

use it for the next six years (Shaw & Barry, 2001). Ford has always denied the Pinto is unsafe compared to other cars of its type and era. The company also argues the Pinto met or surpassed the government's own standards in every model year. However, they neglect to mention that successful lobbying by them and industry associates was responsible for delaying the adoption of any NHTSA crash standard for seven years (Shaw & Barry, 2001).

There are many instances in which groups and organizations have made choices that, from an outside observer's perspective, were easily seen as unethical. Companies such as Enron, British Petroleum, Volkswagen, and a number of major banks and political groups have made decisions that were in their short-term interests while knowing their behavior was either misleading or, in some cases, overtly harmful to at least some of their constituents. Often, the leaders of such organizations are seen as the culprits and are blamed, sued, and sometimes indicted and convicted. However, the major decisions made by companies are rarely attributable to a sole individual. It is often the cooperation of corporate boards and management teams that make most of the decisions for an organization. Therefore, it is quite likely unethical behavior and deception by organizations is at least partially a function of unethical decisions and deception made by groups within the organization.

There is now a fair amount of research on unethical behavior (e.g., deception) in and by groups (for a review, see Messick, 2006). Even in situations where individuals behave cooperatively and abide by prior agreements, groups often defect (i.e., choose a noncooperative response/break agreement to cooperate) from such agreements in order to protect or enhance the group (Morgan & Tindale, 2002; Wildschut, Pinter, Vevea, Insko, & Schopler, 2003). Thus, groups are likely to use the group's welfare to guide their "moral compass" and behave in ways consistent with their self-interest even when it violates typical norms of ethics (Cohen, Gunia, Kim-Jun, & Murnighan, 2009). This "group morality" (Wildschut & Insko, 2006) or *group enhancement/protection norm* (Tindale, 2008), at times, seems to guide group behavior in directions opposite those typically found for individuals. Groups often exacerbate tendencies found for individuals (i.e., group polarization, Kameda, Tindale, & Davis, 2003; Stasser, Kerr, & Davis, 1989), so this discontinuity (Wildschut et al., 2003) is somewhat unique in research on groups and has proved very difficult to change (though see Pinter et al., 2007).

Moreover, there are general differences between individuals and groups in deception use. Research shows deception can be beneficial when negotiating, and groups tend to use deception to their benefit more than individuals (Cohen et al., 2009; Sutter, 2009). Furthermore, under certain circumstances, groups strategically use honesty to maximize their outcomes. However, other research shows lying is more pronounced under team incentives than individual piece-rates (Conrads, Irlenbusch, Rilke, & Walkowitz, 2013). We discuss explanations for these effects and situations where groups

would be more versus less likely to use deception. Finally, based on concepts of social identity theory and ingroup bias (Hogg & Abrams, 1988) and work on group decision making (De Dreu, Nijstad, & van Knippenberg, 2008), we provide a framework for understanding when and why groups use deception.

The ideas underlying the current chapter are drawn from a number of theoretical perspectives and empirical findings. Much of the literature reviewed in this chapter is based on laboratory-based experiments and research involving economic games. Although such research can, at times, be criticized as being artificial, these studies provide compelling empirical evidence for several reasons. First, these experiments help to eliminate individual differences by randomly assigning participants to conditions. Second, using a true experiment helps to eliminate other kinds of confounds as well. Third, these studies capitalize on situational control afforded by the laboratory, which can reduce noise in the outcome measures. Fourth, laboratory studies allow researchers to systematically blend multiple independent variables in order to see how they work together (i.e., statistical interactions) to determine behavior. A final strength of laboratory-based studies is that they provide a unique ability to minimize noise in measurement. Despite these many benefits, there are nevertheless some drawbacks in the use of the experimental design. As mentioned, the biggest issue relates to artificiality. To counter this limitation, researchers often try to increase the experimental realism (i.e., the desired psychological state) of the study.

Therefore, the current chapter organizes the group deception literature by first comparing individuals and groups in mixed-motive situations. There is a discontinuity between individual and group responses to games (e.g., prisoner's dilemma): individuals tend to cooperate while groups tend to compete (Wildschut et al., 2003). In terms of deception, this is interesting as both individuals and groups initially agree to cooperate. We discuss explanations for the effect and their relation to why groups deceive. We then discuss the general differences between individuals and groups in deception use and conclude with a framework for understanding when and why groups use deception.

Mixed-Motive Situations

Groups, by their very nature, have a normative tendency to behave in ways that benefit the group. This has been referred to as the *group enhancement/ protection norm* (Tindale, 2008), in that groups act to both enhance their well-being (e.g., status, wealth) and protect themselves from threats outside the group (often from other groups). This norm is likely a function of evolutionary adaptive pressures associated with the fact that humans live within group contexts (Brewer & Caporael, 2006; Kameda & Tindale, 2006). Because human survival depended on groups remaining together to hunt and fend off predators, groups that could induce members to work toward

enhancing and protecting the group were more likely to survive, as were their members. These tendencies are still present in groups today and can be beneficial in many contexts (e.g., communities pulling together to share resources after a disaster). However, there are situations where the good of the group is not good for non-members or for society at large. For example, company executives may see lying to shareholders so they do not remove their investments from a company as necessary for the company's survival, but it is neither good for the shareholders or for other non-company agents (e.g., clients). In a general sense, deception such as this would be seen as unethical. However, from the perspective of the company executives, the behavior may be seen as necessary for survival and thus acceptable. Below, we discuss research and theoretical reasons for why groups would be more likely to behave unethically in these types of situations than would single individuals.

Perhaps the most well-known demonstration of groups being more likely to engage in self-protection compared to individuals concerns the interindividual-intergroup discontinuity effect (Schopler & Insko, 1992; Wildschut et al., 2003), a well-replicated finding in the small group literature. The basic finding shows a discontinuity between interindividual and intergroup exchanges in mixed-motive situations (i.e., a situation in which an individual or group is tempted to either cooperate or compete). That is, when two individuals communicate while making choices in a prisoner's dilemma game (i.e., a situation in which two people each have two options whose outcome depends on the simultaneous choice made by the other person), they typically agree to cooperate and then subsequently do so when making their individual choices. However, when two small groups play the same game, they agree to cooperate during communication but then typically defect when making their choices. Thus, the discontinuity arises from the finding that rather than exacerbate the dominant individual tendency toward cooperation, groups move in a direction opposite of the individual tendency. This finding conflicts with two well substantiated and related group phenomena: majority factions tend to win and groups tend to polarize.

A number of different explanations for this phenomenon have been proposed (for a review, see Wildschut & Insko, 2007), but research suggests that there are both group-level and intergroup-level aspects to the basic effect. Morgan and Tindale (2002) did focused analysis of the group processes involved in the discontinuity effect. Using a mixed-motive game, they had groups play against other groups or had groups play against individuals. They found the discontinuity effect was stronger when a group was playing against another group (see also Wildschut, Insko, & Pinter, 2007) but was still present when groups played against individuals. When playing against another group, groups often fear being taken advantage of so they defect (i.e., compete) to protect the group. However, when they play against individuals, they no longer fear being taken advantage of but more often use greed as their justification for defection. Morgan and Tindale's (2002) second major finding

was that simply being in a group tended to lead to greater individual preferences among the group members for defection. However, the change was rather small and the majority faction within most groups still typically favored cooperation. Thus, a standard majority wins model would predict groups to be generally cooperative. In contrast, they found that factions favoring defection were more influential than factions favoring cooperation regardless of faction size. In groups where one member preferred defection and two members preferred cooperation prior to group discussion (i.e., the majority faction preferred cooperation), the majority faction only won 33% of the time. Majority factions preferring defection won 88% of the time. Thus, factions that preferred defection were far more influential than factions that preferred cooperation.

Further evidence for the *enhancement/protection norm* is supported by group reactions to dishonesty. In two experiments, Keck (2014) showed dishonesty was punished more often by groups than by individuals, that groups' higher willingness to punish dishonesty was mediated by stronger negative affect, and that increased negative affect in groups is driven by exposure to other group members' negative feelings and opinions during group discussions. Keck (2014) randomly assigned participants to make a decision as a three-person group or an individual using a modified version of the deception game (Gneezy, 2005). The deception game is a decision-making task that requires one party (i.e., the sender) to send a truthful or deceptive message to another party (i.e., the receiver). Specifically, the sender learns of two payment options and is asked to send either a truthful or a deceptive message about the options to the receiver. Sending the truthful message potentially harms the sender's financial outcomes, whereas sending the deceptive message makes it likely the sender will benefit financially (e.g., earn $6 instead of $5). After receiving one of the messages, the receiver ostensibly chooses one of the two payment options based on the sender's message. Thus, the only information the receiver receives about the payoffs is the information included in the sender's message.

Since prior work demonstrated groups are more likely to send deceptive messages than individuals in the deception game (Cohen et al., 2009; Sutter, 2009), Keck (2014) focused on reactions to deceptive messages. Thus, all participants were assigned to the role of receiver and all senders were actually computers. In the game, receivers tried to guess a random number between 0 and 1000. Those who answered correctly earned €7.50 and all others received €3.50. Participants were told they were paired with another player (i.e., the sender) who knew the correct value of the number and would send them a message before they made their guess. Half of the participants were told the sender was another individual while the other half were told the sender was a group of three individuals. Senders were described as having an incentive to lie because a wrong guess by the receiver would result in a higher payoff for the sender (€7.50 versus €3.50). However, the sender's message

was the only information the receivers were given to make their decision. In actuality, all participants were sent a deceptive message containing the same wrong number. After making their decision, participants were informed of the correct number.

Afterwards, participants were told the game would continue with a second stage. Participants were told at the beginning of the game that they would progress to the second stage or end after the first stage based on random assignment from a computer. In actuality, all participants progressed to the second stage. During the second stage, participants were given the option of spending some of their payoffs on punishing senders for sending them an incorrect number. Participants could spend between €0.10 and €1.00 (in 10-cent intervals) to lower senders' payoffs by four times the amount spent. Keck (2014) showed there were higher levels of mutually harmful spending on punishment when groups made punishment decisions. The effect was mediated by the stronger degree of negative affect that group members, relative to individuals, felt when interacting with a dishonest party. Results also showed diffusion of responsibility did not function as an alternative mediator. There was no evidence groups (compared to individuals) focused more on maximizing their financial self-interest when deciding whether to punish dishonest behavior. Also, willingness to punish dishonesty did not depend on the target of the punishment. Groups were punished just as much as individuals. This suggests group members' greater desire to punish dishonesty could be attributed to factors specific to group decision making rather than factors related to the source of the dishonest behavior.

Keck (2014) used the same procedure in a second study, but in order to focus on the factors that were driving the negative affect and punishment in groups, Keck measured negative affect twice in the group decision conditions: before (Time 1) and after (Time 2) the group discussion took place. The results again showed groups were more willing than individuals to punish dishonest behavior even if punishment was financially costly. As in Study 1, the effect was mediated by greater negative affect in groups compared to individuals. The results also showed although there was no difference in negative affect between individuals and groups before the group discussion, group members reported significantly more negative affect than individuals after talking to each other. Thus, these results provide evidence that the heightened negative affect in groups was caused by the interaction among group members. Taken together, Keck's (2014) results suggest being part of a group increases negative emotions toward dishonest others and leads to a greater willingness to utilize costly punishment in order to protect the ingroup.

Although there is evidence that groups will behave uncooperatively for strategic reasons (see Bornstein & Yaniv, 1998), recent research has found groups still choose defection in economic games where it is not the dominant response. Tindale et al. (2006) found similar asymmetries in social influence

patterns (i.e., minority factions within a group favoring defection winning out over majority factions favoring cooperation) for groups playing games in which the mutual defection response was the worst option possible and the difference in payoff for cooperate/cooperate or cooperate/defect response combinations was very slight. The study involved multiple plays of the game and a single defection by either team at any point during the experiment tended to lead to mutual defection for all subsequent plays. Thus, groups tended to defect even when it was not economically rational to do so.

Additional research has also shown group behavior in these settings is generally driven by concerns for the welfare of the ingroup rather than attempts to disadvantage the outgroup (Halvey, Bornstein, & Sagiv, 2008). Halvey et al. (2008) gave groups options to either cooperate, defect with a benefit to the ingroup but no additional loss to the outgroup, or defect with a penalty to the outgroup but no benefit to the ingroup in an intergroup mixed-motive game. In all cases, groups chose the benefit to the ingroup choice. Thus, group behavior in these settings seems to be driven by motives to either protect or enhance (or both) the ingroup.

Two main perspectives describe the reasons for differences in the behavior of groups and individuals in mixed-motive situations (for reviews, see Cohen, Meier, Hinsz, & Insko, 2010; Wildschut & Insko, 2007). According to the fear-and-greed explanation, fear and greed characterize intergroup interactions more than they characterize interpersonal interactions (Cohen et al., 2010; Wildschut & Insko, 2007; Wildschut et al., 2003). This explanation assumes groups are more likely than individuals to be distrusted (i.e., people fear groups), and groups are more likely to attempt to maximize their own outcomes, either in an absolute or relative sense (i.e., groups are greedier than individuals). Groups, more than individuals, tend to be fearful of being taken advantage of by the other group. However, even when playing the game against a single individual, groups still are more likely to choose non-cooperation, thinking they can take advantage of the more cooperative individual (Morgan & Tindale, 2002). Thus, groups both protect themselves by choosing non-cooperation, but also attempt to ensure that they do as well or better than the other player in the situation. Interestingly, there is little evidence that the effect stems from wanting to hurt the outgroup.

Morgan and Tindale (2002) showed this effect is at least partly due to asymmetries in the influence processes among the group members. Prior to making a group choice as to whether to cooperate or not, they asked each member to privately note their individual preference. Although preferences for cooperation were slightly lower when playing against a group, most individual members favored cooperation regardless of the type of opponent. Thus, most of the groups entered the discussion with majority factions favoring cooperation. However, minority factions favoring non-cooperation were quite persuasive and won out over cooperative majority factions two-thirds of the time, leading the majority factions to adopt a non-cooperative

response. Majority factions favoring non-cooperation virtually never lost to minority factions favoring cooperation. Tindale (2008) argued the shared motivation to protect or enhance the group (i.e., the *group enhancement/protection norm*) acted much like other *shared task representations* (described in more detail below) and made the non-cooperative response easier to defend because it was consistent with enhancing and protecting the ingroup.

Further evidence supporting the fear-and-greed explanation comes from studies using the PDG-Alt (Insko, Schopler, Hoyle, Dardis, & Graetz, 1990). The PDG-Alt is a variation of the prisoner's dilemma game described above; however, in this version of the game, a third choice (i.e., withdrawal) is added that guarantees equal intermediate outcomes (i.e., outcomes intermediate to those obtained when both players cooperate or both players compete) for both sides. In the PDG-Alt, competition is evidence of self-interest or greed whereas withdrawal is evidence of distrust or fear. In the PDG-Alt, groups compete more, withdraw more, and cooperate less than individuals (Insko et al., 1990).

A second explanation for why group and individual behavior differ in mixed-motive situations is groups are better at problem-solving than individuals (Bornstein, Kugler, & Ziegelmeyer, 2004; Lodewijkx, Rabbie, & Visser, 2006; Thompson, Peterson, & Brodt, 1996). According to the group-decision making explanation, "two heads are better than one" when it comes to solving complex economic problems. Bornstein et al. (2004; Bornstein & Yaniv, 1998) have provided evidence consistent with this explanation by showing that groups behave more consistently with game-theoretic predictions in economic games. Likewise, Thompson et al. (1996) found groups were better than individuals at achieving Pareto efficient outcomes (i.e., a state in which it is impossible to reallocate resources so as to make any one individual or preference criterion better off without making at least one individual or preference criterion worse off) in a multi-issue negotiation. However, because game-theoretic rationality and greed both involve self-interested behavior, there is some debate as to whether groups are actually more rational than individuals or whether they are simply more focused on winning or not losing (c.f., Bornstein et al., 2004; Lodewijkx et al., 2006; Wildschut & Insko, 2007).

INDIVIDUAL AND GROUP DECEPTION DIFFERENCES

Although many studies have compared intergroup and interpersonal interactions in mixed-motive economic games (e.g., Bornstein et al., 2004; Hargreaves & Zizzo, 2009; Wildschut et al., 2003), research comparing group and individual deception is scarce. In one of the few studies that investigates this topic, Sutter (2009) examined group and individual lying with the deception game (Gneezy, 2005). As previously described, the deception game is an economic decision-making task that requires one party (i.e., the sender)

to send a truthful or deceptive message to another party (i.e., the receiver), and sending the truthful message potentially harms the sender's financial outcomes whereas sending the deceptive message makes it likely the sender will benefit financially (e.g., earn $6 instead of $5).

Research with the deception game has found 36% of university students sampled lie (Gneezy, 2005) and men lie more than women (55% and 38%, respectively; Dreber & Johannesson, 2008). Sutter (2009) found groups lied less than individuals (23% and 44%, respectively) but suggested this was a function of groups expecting to be distrusted (i.e., groups told the truth only because they expected their message to be disbelieved). Sutter's (2009) results are consistent with the fear-and-greed explanation of the discontinuity effect in that groups expected to be distrusted (i.e., fear) and they acted strategically to maximize their outcomes (i.e., greed). Sutter's (2009) findings suggest in certain circumstances, groups strategically use honesty to maximize their outcomes. However, as mentioned above, previous work has found groups choose uncooperative responses in economic games where such responses are not the dominant response (Bornstein & Yaniv, 1998; Tindale et al., 2006). Therefore, it seems motives to enhance or protect (or both) the ingroup at least partially explain Sutter's (2009) results.

Cohen et al. (2009) extended Sutter's (2009) work by testing whether groups are more deceptive than individuals when lying yields a higher payoff than honesty does. They used a modified version of the deception game (Gneezy, 2005), in which all of the participants sent a computer-mediated message about the payoffs to anonymous receivers (who did not exist). These receivers ostensibly chose between two payment options (each gave $5 to one party and $6 to the other) but they did not know which option gave them the higher payoff. Supposedly, the receivers would use the participants' message to guide their choice. Group payoffs were $15 and $18 respectively; they were required to divide the money equally. After talking about (in the group conditions) or thinking about (in the individual conditions) their message choice for three minutes, participants either told the truth or lied about the payoffs. In the study, the uncertain condition was the standard deception game, in which participants did not know whether the receiver would believe their message (Gneezy, 2005; Sutter, 2009). In the certain condition, participants learned that receivers had preemptively committed to following their payoff-choice recommendation. Thus, participants in the certain condition knew that receivers would choose the option they identified as giving receivers more money. These instructions made it clear that deception was guaranteed to give each participant $6 and honesty was guaranteed to give each participant $5.

Cohen et al. (2009) found groups lied more than individuals when the receiver's response was certain, but groups lied relatively less than individuals when the receiver's response was uncertain. Specifically, when the message was certain to be followed, almost half of the individuals lied (48%)

but 82% of groups lied. Consistent with prior deception game studies (Gneezy, 2005; Sutter, 2009), when it was uncertain whether the message would be followed, 32% of individuals lied whereas 24% of groups lied. Not only did groups lie more when they were sure that they would be believed, they also reported more self-interest. Most groups who lied reported no qualms about using deception. Further analyses indicated that self-interest explained why groups lied more than individuals when the receiver's response was certain. Previous research suggests groups are greedier than individuals and their greed fuels competition (Insko et al., 1990; Wildschut & Insko, 2007). Cohen et al. (2009) results suggest greed also seems to fuel deception.

Looking at the influence of compensation schemes on deceptive behavior, Conrads et al. (2013) employed a simple experimental design introduced by Fischbacher and Heusi (2008). Researchers compared the incentives to lie under two schemes: a team compensation scheme and an individual piece-rate scheme. Results revealed under the team compensation scheme (i.e., the random production output of two agents was pooled and each agent received one half of a compensation unit for each unit of the joint production output), the marginal gain from lying (i.e., the return from exaggerating the own production output by one unit) was about half of the gain under the individual piece-rate scheme (i.e., for each unit of random production output, the agent received one compensation unit). These findings suggest lying is more pronounced under the individual piece-rate scheme than under the team incentive scheme. However, lying under the team incentive scheme is not exclusively beneficial for oneself, contrary to the individual compensation scheme. It also benefits the other agent in the team. Thus, an agent under a team incentive scheme, as opposed to an individual scheme, might be more able to justify such a lie. Indeed, this latter interpretation is also consistent with the *group enhancement/protection norm* (Tindale, 2008).

Supporting this idea, Conrads et al. (2013) found lying was prevalent under both team incentives and individual piece-rate compensation schemes, but the effect was more pronounced under team incentives. This indicates groups and organizations are well advised to be vigilant regarding potentially harmful side effects of compensation schemes. Those working under team incentives might be particularly prone to lying and deception because they might be able to more easily justify that lying led to a positive outcome (e.g., benefit other team members).

Deception driven by the motivation to enhance and/or protect the ingroup is a driving force behind organizational scandals. The fear and greed of groups seems to predispose them to lie more than individuals (c.f., Tindale, Smith, Thomas, Filkins, & Sheffey, 1996; Wildschut & Insko, 2007). In some situations, having groups make decisions may be particularly risky when organizations anticipate tradeoffs between ethics and self-interest. However, groups can also appear to be exemplars of honesty when there is a sense in the group that it is defined as being honest and trustworthy

(e.g., Cohen et al., 2009; Stawiski, Tindale, & Dykema-Engblade, 2009). Taken together, research suggests individuals and groups view honesty and deception differently. In particular, an ethical issue for individuals may be a strategic or normative issue for groups. Thus, whether groups will handle sensitive situations in an honest manner may depend on the group's preference for ethics or economics in that particular situation. However, these two preferences need not be mutually exclusive.

Framework

To understand when and why groups use deception, we need to examine how groups process information. Group cognition consists of more than members simply sharing specific pieces of information and preferences (Resnick, Levine, & Teasley, 1991; Tindale & Kameda, 2000). Laughlin (1980, 2011) has argued one of the reasons groups are better problem solvers than individuals is group members often share a conceptual system that allows them to realize when a proposed solution is correct within that system. This shared conceptual system, or background knowledge, is what allows a minority faction with a correct answer to influence a larger incorrect faction to change its preference to the correct alternative. For example, suppose there is a group of five people discussing what 4^2 is equal to. If all five group members understand the principles of arithmetic, then the group shares a conceptual system (i.e., they all have background knowledge about arithmetic they can apply to their current discussion).

These situations are well described by social decision scheme models, called truth wins and truth supported wins (Laughlin, 1980). The truth wins model predicts any group that has at least one member with the objectively correct answer will be able to solve the problem correctly (Laughlin, 1980). The truth supported wins model argues at least two members of the group must have the correct answer in order for the group to solve the problem correctly (Laughlin, 1980). For groups with more than four members, both models predict minority influence for minority factions with the objectively correct answer. Laughlin and Ellis (1986) proposed such minority influence processes are likely to occur for demonstrable or intellective tasks (i.e., those that have a demonstrably correct solution) and the shared conceptual system is a key component of demonstrability. For judgmental tasks (i.e., those without a demonstrably correct solution), majority/plurality processes are more likely to occur.

Returning to the arithmetic example, even if four of the five group members believe 4^2 is equal to 8, as long as one of the five group members knows that 4^2 is equal to 16, then the group should arrive at 16 for their final decision. This is what the truth wins model predicts (i.e., only one person in the group needs to have the objectively correct answer). However, the truth supported wins model argues at least two of the five need to know 4^2 is equal

to 16 for the group to produce the correct answer. Also, since there are more than four members in this hypothetical group, both models predict the minority faction should be able to convince the majority faction they have the correct answer. This minority influence is much more likely for a demonstrable task such as the current example (i.e., what is 4^2 equal to) because it does have an objectively correct answer. If, for instance, the group was dealing with a judgmental task (e.g., determining guilt), the group would more likely endorse the majority faction's position because there is no objectively correct answer and the group's shared conceptual system is weaker.

Tindale et al. (1996) argued the shared conceptual system underlying demonstrability is one instance of a *shared task representation*. They defined a *shared task representation* as "any task/situation relevant concept, norm, perspective, or cognitive process that is shared by most or all of the group members" (Tindale et al., 1996, p. 84). Task/situation relevant means the representation must have implications for the choice alternatives involved, and the degree to which a shared representation affects group decision processes and outcomes will vary as a function of its relevance. Its influence will also vary by the degree to which it is shared among the group members: the greater the degree of sharedness (i.e., the more members who share it), the greater its influence. If no *shared task representation* exists, or if multiple conflicting representations are present, groups will tend to follow a symmetric majority/plurality process. However, when a s*hared task representation* does exist, the group decision process will tend to become asymmetric in favor of alternatives that fit within or are supported by the representation. Under such conditions, majorities/pluralities favoring an alternative consistent with the shared representation are more powerful than are identically sized majorities/pluralities favoring alternatives that are not consistent with or supported by the representation. In addition, minority factions favoring an alternative consistent with the shared representation can sometimes be more influential than majority factions favoring an alternative inconsistent with the shared representation, even when the majority is changing to a logically or normatively incorrect position (e.g., ignoring base rates).

A number of theories are consistent with or can explain the presence of an ingroup *enhancement/protection norm*. Social identity theory and self-categorization theory argue group identification leads directly to ingroup favoritism and other behaviors that differentiate one's group from others (Hogg & Abrams, 1988; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Work on the role of groups in evolutionary adaptation of the species argues living and hunting in groups had survival implications and being rejected by the group could lead to devastating outcomes such as starvation and death (Brewer & Caporael, 2006; Levine & Kerr, 2007). Additional simulation studies (Choi & Bowles, 2007) have found societies with a substantial number of members who are parochial altruists (i.e., those who sacrifice for the ingroup and shun or aggress against outgroup members) tend to be stable while societies

with mainly non-exclusive altruists die off over time. More recent work has begun to isolate the physiological and neurological correlates of these effects and has shown that oxytocin helps to regulate responses to both ingroup and outgroup members (De Dreu et al., 2010). Thus, behaving in ways that favor the group welfare appears to be adaptive for both the group and the individuals that depend on it. Consequently, the *group enhancement/protection norm* is probably well ingrained in most group settings. Once group members begin to think of themselves as a group, they will begin to favor options that protect or enhance the group welfare. Thus, the *group enhancement/protection norm* serves as a *shared task representation* in that members share the norm (even if they do not explicitly recognize it), and behavioral options that are consistent with the norm will be more likely to be adopted than options inconsistent with the norm.

In many situations, such behavior will have few if any implications for people outside the group and may even be perceived as ethical both within and outside the group (e.g., helping a neighbor or family member, working extra hours to help insure the company does well this quarter). However, there are situations where group-serving behavior has negative consequences for the parties associated with the behavior and potentially society as a whole. For example, during the financial crisis of 2007–2008, predatory lending practices were very profitable for banks, and such profits were probably the motivating factor underlying these lending practices. It is unlikely the banks were motivated to hurt the borrowers or create havoc in the economy. Yet, such behavior did, in fact, have such consequences and those consequences were foreseeable. Thus, many groups will not be prone to act unethically in all or even most situations, but they will often choose alternatives that are in their best interest, even when non-group members might perceive those choices as unethical.

So, if groups' natural tendencies are to behave in ways that support the groups' welfare and status, even when doing so leads to unethical behavior, how can we get groups to go against their nature and behave ethically? Work by Cohen et al. (2009) has shown making honesty a strategic choice that benefits the group will lead groups to be just as honest as individuals if not more so. However, changing strategic interests so ethical behavior leads to the best economic outcomes may not always be plausible. Another possible strategy is to make groups feel like their best or "true" interests are associated with ethical responses. In other words, one can try to change what the group (or its members) sees as "their nature" or "in their best interest." Research on social identity has shown that when group membership is salient (e.g., referencing an outgroup or increasing accountability to the ingroup), members' behavior tends to conform to what the members see as the group norm (Hogg & Abrams, 1988; Postmes & Spears, 1998). Thus, if a group member can create a sense that the group is defined as being honest and trustworthy, then such behaviors would be normatively correct within the group

and may serve as a *shared representation* for tasks that involve ethical aspects. In other words, if the group members define the group as honest and trustworthy, the *group enhancement/protection norm* should encourage honest and trustworthy behavior because maintaining that positive identity is in the group's best interest. For example, consider a community that has suffered significant damage after a natural disaster. If the *group enhancement/protection norm* defines a trustworthy and cooperative group, the community will be much more likely to pull together and share resources after the disaster. However, if the *group enhancement/protection norm* defines a deceitful and non-cooperative group, the community will be unlikely to pull together and community members may lie to one another about the resources they actually have. Under such circumstances, groups may be more likely to behave in ways that protect and enhance the honesty and trustworthiness of the group and avoid the less ethical direction implied by economic issues.

De Dreu et al. (2008) developed a model of group judgment and decision making based on the combination of epistemic and social motives. Called the motivated information processing in groups (MIP-G) model, the model argues information processing in groups is better understood by incorporating two somewhat orthogonal motives: (1) high versus low epistemic motivation (i.e., willingness to expend effort to achieve a thorough, rich, and accurate understanding of the world) and (2) prosocial (i.e., concerned with joint outcomes and fairness) versus proself (i.e., concerned with own outcomes) motivation. Earlier work on negotiation had shown that negotiators who shared high epistemic and prosocial motivations were better able to find mutually beneficial tradeoffs and reach better integrative agreements as compared to negotiators with any other combination of motives (De Dreu, 2010). Research now suggests the same appears to be true for groups working cooperatively to solve a problem or make a decision. According to the MIP-G model, high epistemic motivation involves a goal to be accurate or correct, which should lead to deeper and more thorough information search and analysis (Kruglanski & Webster, 1996). Work on these information sharing effects has consistently demonstrated that instilling a goal of accuracy or defining the task in terms of solving a problem both increase information sharing (Postmes, Spears, & Cihangir, 2001; Stewart & Stasser, 1995). Members high in prosocial motivation help to insure that all types of information held by each member are likely to be disseminated, rather than just information supporting the position held by an individual member. Consistent with this assertion, other research has shown group members who focus on preferences rather than information tend to impede information sharing (Mojzisch & Schutz-Hardt, 2010).

According to MIP-G, proself group members are less likely to input information conducive to group goals and collective functioning, and they are less likely to disseminate information in an accurate way compared to prosocial group members. Proself group members are also more likely to spin

information conducive to personal goals and preferences, to strategically withhold information, and to engage in lying and deception compared to prosocial group members (De Dreu et al., 2008). These tendencies should amplify when epistemic motivation among group members is high rather than low. For example, someone who is willing to expend effort to achieve a thorough, rich, and accurate understanding of the group situation (i.e., high epistemic motivation) is likely to seek out information until the group member is able to make sense of the situation. However, proself motivation biases the type of information the member looks for, generates, and processes. Thus, the group member is more likely to be concerned with self-interests and to ignore other group members' needs, interests, and beliefs (De Dreu et al., 2008).

This is because higher levels of epistemic motivation create a stronger tendency to deliberately and systematically process the information during group discussion. Also, higher levels of epistemic motivation reduce tendencies toward group centeredness and concomitant preference for autocratic leadership and reduced participative decision making. Thus, the MIP-G model predicts group information processing will only approach optimal levels when group members are high on both epistemic motivation and prosocial orientation. This is because high epistemic motivation and prosocial orientation is the only combination that produces both systematic and thorough processing of information in an unbiased manner. Although the model is fairly recent, it does a good job of explaining a number of well-replicated findings and has fared well in the few direct attempts to test it (Bechtoldt, De Dreu, Nijstad, & Choi, 2010; De Dreu, 2007).

IMPLICATIONS AND CONCLUSION

In summary, we have argued, based on a large amount of empirical evidence, groups are naturally prone to behave in ways that enhance and/or protect the group. When people see themselves as part of a group, their responses become normative and form a framework within which members interpret their behaviors (i.e., a *shared task representation*). Thus, even if these group-normative responses are not initially favored by all of the group members, those members favoring such responses will tend to be quite persuasive. This tendency will not always make groups more unethical or deceptive than individuals, and in some cases, it may actually make groups more ethical and honest. However, whenever a group is making a decision that has implications for the welfare of the group, choice alternatives that enhance or protect the group welfare become easier to defend in the group discussion and will often be chosen by the group. Even in situations where an outside observer might define such responses as unethical (e.g., lying in a negotiation, failing to disclose relevant information, breaking former agreements), groups will still be prone to perform such unethical responses because these responses

are perceived as normative and good for the group. However, other types of behaviors can also be seen as normative or "good for the group" depending on how the group defines itself. By changing how the group defines or views itself, it is possible to make other, more ethical responses appear normative and best for the group, and ultimately move the group in a more ethical direction.

While the empirical evidence does provide insight as to whether and when groups will be more likely to use deception than individuals, future work is certainly needed to further our understanding of deception within group contexts. In this chapter, we have proposed the *group enhancement/protection norm* serves as a *shared task representation* in that members share the norm (even if they do not explicitly recognize it) and behavioral options that are consistent with the norm will be more likely to be adopted than options inconsistent with the norm. Additionally, we outlined De Dreu et al.'s (2008) MIP-G model as another framework by which to conceptualize group deception use. We want to note that these frameworks are not necessarily mutually exclusive. It is completely conceivable that a group could hold a *shared task representation* that signals high epistemic and high prosocial motivations. In such a conception, both frameworks would predict reduced deception and increased ethical behavior by groups. However, would these frameworks make the same predictions in both within and between group situations? That is, could high epistemic and high prosocial motivation lead to ethical decisions for within group situations but unethical decisions between groups due to the group protection/enhancement norm? How might altering the group's *shared task representation* influence these motivations and in turn group deception use? Future research should aim to dissect and clarify these and other relationships.

Many of today's most serious issues revolve around notions of ethics and how group membership can alter or exacerbate unethical tendencies in groups. From terrorism, to financial crises, to the political climate, a number of group-centric or group-serving ideas have been used to promote behavior that can often be perceived as unethical outside of the specific group context. The research covered in this chapter attempts to further understand the group-level variables that affect unethical and deceptive behavior and shows how groups might be able to use these same processes to attenuate or prevent deception. Research on mixed-motive situations has shown groups tend to behave unethically and lie more often compared to individuals in the same situation. The fear and greed explanation argues groups are more likely than individuals to be distrusted (i.e., people fear groups), and groups are more likely to attempt to maximize their own outcomes, either in an absolute or relative sense (i.e., groups are greedier than individuals). This is also consistent with the idea that group members will behave in ways consistent with *shared task representations* such as the *group enhancement/protection norm*. Research in this area has also shown group behavior in these settings is generally driven

by concerns for the welfare of the ingroup rather than attempts to disadvantage the outgroup. Taken together, research points to the tendency for groups to lie and deceive more often than individuals because doing so often protects and enhances the ingroup's welfare. However, when being honest strategically benefits the group, the group will likely be just as honest—if not more so—than individuals in similar situations.

The potential benefits to society are vast if groups could be moved to behave more ethically by changing the ways group members perceive or think of the group or the task at hand. Because important decisions are often made by groups, and group-serving perspectives are commonly salient in such situations, obtaining a better understanding of how such perspectives affect groups and how they may be altered to enhance ethical concerns should prove valuable in numerous decision-making contexts. Team and organizational leaders can use the knowledge generated from the literature discussed to implement strategies for facilitating more ethical decision making in their own groups. For instance, team leaders could be trained on strategies for promoting ethical norms within their groups and to help teams adopt a promotion mindset (i.e., an emphasis toward progress, advancement, and gaining rather than maintaining the status quo) once these norms are firmly established.

The theoretical perspectives and empirical findings covered here may also aid in designing educational and training materials that increase the role of ethics in decision making. We hope groups and organizations will be able to use the information from the current chapter to enhance the role of ethics in leadership training. Similar techniques may be useful for designing role-playing exercises for business ethics courses and could be modified as exercises for critical thinking courses in high schools and colleges. This information may also help to inform policy issues associated with unethical behavior by groups in other domains (e.g., terrorist groups, gangs, juries).

REFERENCES

Bechtoldt, M. N., De Dreu, C. K., Nijstad, B. A., & Choi, H. S. (2010). Motivated information processing, social tuning, and group creativity. *Journal of Personality and Social Psychology, 99*(4), 622.

Bornstein, G., Kugler, T., & Ziegelmeyer, A. (2004). Individual and group decisions in the centipede game: Are groups more "rational" players? *Journal of Experimental Social Psychology, 40*(5), 599–605.

Bornstein, G., & Yaniv, I. (1998). Individual and group behavior in the ultimatum game: Are groups more "rational" players? *Experimental Economics, 1*(1), 101–108.

Brewer, M. B., & Caporael, L. R. (2006). An evolutionary perspective on social identity: Revisiting groups. In M. Schaller, J. A. Simpson, & D. T. Kenrick (Eds.), *Evolution and Social Psychology* (pp. 143–161). Madison, CT: Psychosocial Press.

Choi, J. K., & Bowles, S. (2007). The coevolution of parochial altruism and war. *Science, 318*(5850), 636–664.

Cohen, T. R., Gunia, B. C., Kim-Jun, S. Y., & Murnighan, J. K. (2009). Do groups lie more than individuals? Honesty and deception as a function of strategic self-interest. *Journal of Experimental Social Psychology, 45*(6), 1321–1324.

Cohen, T. R., Meier, B. P., Hinsz, V. B., & Insko, C. A. (2010). Competitive group interactions: Why they exist and how to overcome them. In S. Schuman (Ed.), *The handbook for working with difficult groups* (pp. 223–236). San Francisco: Jossey-Bass.

Conrads, J., Irlenbusch, B., Rilke, R. M., & Walkowitz, G. (2013). Lying and team incentives. *Journal of Economic Psychology, 34,* 1–7.

De Dreu, C. (2007). Cooperative outcome interdependence, task reflexivity, and team effectiveness: A motivated information processing perspective. *Journal of Applied Psychology, 92*(3), 628–638.

De Deru, C. K. (2010). Social value orientation moderates ingroup love but not outgroup hate in competitive intergroup conflict. *Group Processes & Intergroup Relations, 13*(6), 701–713.

De Dreu, C. K., Greer, L. L., Handgraaf, M. J., Shalvi, S., Van Kleef, G. A., Baas, M., …, Feith, S. W. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science, 328*(5984), 1408–1411.

De Dreu, C. K., Nijstad, B. A., & van Knippenberg, D. (2008). Motivated information processing in group judgment and decision making. *Personality and Social Psychology Review, 12*(1), 22–49.

Dreber, A., & Johannesson, M. (2008). Gender differences in deception. *Economics Letters, 99*(1), 197–199.

Fischbacher, U., & Heusi, F. (2008). *Lies in disguise: An experimental study on cheating* (Research Paper Series). Thurgau Institute of Economics and Department of Economics at the University of Konstanz.

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review, 95*(1), 384–394.

Halevy, N., Bornstein, G., & Sagiv, L. (2008). "In-group love" and "out-group hate" as motives for individual participation in intergroup conflict: A new game paradigm. *Psychological Science, 19*(4), 405–411.

Hargreaves Heap, S. P., & Zizzo, D. J. (2009). The value of groups. *The American Economic Review, 99*(1), 295–323.

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes.* Florence, KY: Taylor & Frances and Routledge.

Insko, C. A., Schopler, J., Hoyle, R. H., Dardis, G. J., & Graetz, K. A. (1990). Individual-group discontinuity as a function of fear and greed. *Journal of Personality and Social Psychology, 58*(1), 68–79.

Kameda, T., & Tindale, R. S. (2006). Groups as adaptive devices: Human docility and group aggregation mechanisms in evolutionary context. In M. Schaller, J. A. Simpson, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 317–341). New York: Psychology Press.

Kameda, T., Tindale, R. S., & Davis, J. H. (2003). Cognitions, preferences, and social sharedness: Past, present, and future directions in group decision making. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 458–485). New York: Cambridge University Press.

Keck, S. (2014). Group reactions to dishonesty. *Organizational Behavior and Human Decision Processes, 124*(1), 1–10.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review, 103*(2), 263–283.

Laughlin, P. R. (1980). Social combination processes of cooperative problem-solving groups on verbal intellective tasks. *Progress in Social Psychology, 1,* 127–155.

Laughlin, P. R. (2011). *Group problem solving.* Princeton, NJ: Princeton University Press.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*(3), 177–189.

Levine, J. M., & Kerr, N. L. (2007). Inclusion and exclusion: Implications for group processes. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 759–784). New York: Guilford Press.

Lodewijkx, H. F., Rabbie, J. M., & Visser, L. (2006). "Better to be safe than to be sorry": Extinguishing the individual—Group discontinuity effect in competition by cautious reciprocation. *European Review of Social Psychology, 17*(1), 185–232.

Messick, D. M. (2006). Ethics in groups: The road to hell. In E. Mannix, M. Neale, & A. Tenbrunsel (Eds.), *Research on managing groups and teams: Ethics in groups* (Vol. 8). Oxford, UK: Elsevier Science Press.

Mojzisch, A., & Schulz-Hardt, S. (2010). Knowing others' preferences degrades the quality of group decisions. *Journal of Personality and Social Psychology, 98,* 794–808. https://doi.org/10.1037/a0017627.

Morgan, P. M., & Tindale, R. S. (2002). Group vs individual performance in mixed-motive situations: Exploring an inconsistency. *Organizational Behavior and Human Decision Processes, 87*(1), 44–65.

Pinter, B., Insko, C. A., Wildschut, T., Kirchner, J. L., Montoya, R. M., & Wolf, S. T. (2007). Reduction of interindividual-intergroup discontinuity: The role of leader accountability and proneness to guilt. *Journal of Personality and Social Psychology, 93*(2), 250–265.

Postmes, T., & Spears, R. (1998). Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin, 123*(3), 238–259.

Postmes, T., Spears, R., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of Personality and Social Psychology, 80*(6), 918–930.

Resnick, L. B., Levine, J. M., & Teasley, S. D. (1991). *Perspectives on socially shared cognition.* Washington, DC: American Psychological Association.

Schopler, J., & Insko, C. A. (1992). The discontinuity effect in interpersonal and intergroup relations: Generality and mediation. *European Review of Social Psychology, 3*(1), 121–151.

Shaw, W. H., & Barry, V. E. (2001). *Moral issues in business* (8th ed.). Belmont, CA: Wadsworth.

Stasser, G., Kerr, N. L., & Davis, J. H. (1989). Influence processes and consensus models in decision-making groups. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 279–326). Hillsdale, NJ: Lawrence Erlbaum.

Stawiski, S., Tindale, R. S., & Dykema-Engblade, A. (2009). The effects of ethical climate on group and individual level deception in negotiation. *International Journal of Conflict Management, 20*(3), 287–308.

Stewart, D. D., & Stasser, G. (1995). Expert role assignment and information sampling during collective recall and decision making. *Journal of Personality and Social Psychology, 69*(4), 619–628.

Sutter, M. (2009). Individual behavior and group membership: Comment. *The American Economic Review, 99*(5), 2247–2257.

Thompson, L., Peterson, E., & Brodt, S. E. (1996). Team negotiation: An examination of integrative and distributive bargaining. *Journal of Personality and Social Psychology, 70*(3), 66–78.

Tindale, R. S. (2008). *The wisdom (an occasional lack thereof) of groups*. Presidential Address Presented at the Midwestern Psychological Association Annual Convention, Chicago, IL.

Tindale, R. S., & Kameda, T. (2000). 'Social sharedness' as a unifying theme for information processing in groups. *Group Processes & Intergroup Relations, 3*(2), 123–140.

Tindale, R. S., Morgan, P., Stawiski, S., Dykema-Engblade, A., Meisenhelder, H., Wittkowski, E., & Jacobs, E. (2006). *Further explorations of the individual-group discontinuity effect*. Paper Presented at the First Annual INGRoup Conference, Pittsburgh, PA.

Tindale, R. S., Smith, C. M., Thomas, L. S., Filkins, J., & Sheffey, S. (1996). Shared representations and asymmetric social influence processes in small groups. In E. Witte & J. H. Davis (Eds.), *Understanding group behavior: Consensual action by small groups* (pp. 81–103). Mahwah, NJ: Lawrence Erlbaum.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA: Basil Blackwell.

Wildschut, T., & Insko, C. A. (2006). A paradox of individual and group morality: Social psychology as empirical philosophy. In P. A. M. Van Lange (Ed.), *Bridging social psychology: Benefits of transdisciplinary approaches* (pp. 377–384). Hillsdale, NJ: Lawrence Erlbaum.

Wildschut, T., & Insko, C. A. (2007). Explanations of interindividual—Intergroup discontinuity: A review of the evidence. *European Review of Social Psychology, 18*(1), 175–211.

Wildschut, T., Insko, C. A., & Pinter, B. (2007). Interindividual—Intergroup discontinuity as a joint function of acting as a group and interacting with a group. *European Journal of Social Psychology, 37*(2), 390–399.

Wildschut, T., Pinter, B., Vevea, J. L., Insko, C. A., & Schopler, J. (2003). Beyond the group mind: A quantitative review of the interindividual-intergroup discontinuity effect. *Psychological Bulletin, 129*(5), 698–722.

Organizational Deception: Lies at Work

Anne P. Hubbell

Any lie in an organization can be a high-stakes lie. Lies in the workplace can result in an individual being written up, decrease in productivity, loss of trust, fraud, bankruptcy, and even lawsuits. The estimated cost of "dishonesty in business in the United States range[s] from $6 to $200 billion annually" (Edwards & Nadler, 2014, p. 106). Enron, for example, once the seventh largest company in the US, filed bankruptcy, putting thousands out of jobs because they misrepresented the income of the organization ("The Fall of Enron," 2016, para 3). But corporate-level deception is not the only issue in organizations. In a self-report study, 45% of employees from multiple organizations stated that they have lied to someone in their organization (Lindsey, Dunbar, & Russell, 2011). This chapter includes an examination of organizational deception from both the corporate and interpersonal levels. First, corporate-level deception will be examined through discussions of strategic ambiguity (Eisenberg, 1984), fraud, and deception in advertising. Second, suggestions for reducing deception at the corporate level will be considered. Third, research from the interpersonal level of organizations will be explored with a focus on reducing lying in organizations. The final section includes recommendations for future research.

A. P. Hubbell (✉)
Department of Communication Studies,
New Mexico State University, Las Cruces, NM, USA
e-mail: ahubbell@nmsu.edu

Definition(s) and Deception in the Organizational Literature

Definitions of deception in organizations rely on interpersonal deceptive communication literature. Organizational deception researchers focus on the intent of the liar, such that "deception occurs when one individual conveys information to another with the purposeful intent to mislead" (Dunbar et al., 2014, p. 854). Deception research in organizations also often takes a strategically interactional perspective in that "the perpetrator *knows* the information is false, *wants* to mislead the other person, and engages in the behavior proactively" (Grover, 1997, p. 69, emphasis in original). Deception research at both the corporate and individual/interpersonal levels in organizations is interactional in that much of the research focuses on lying as a response to a context, situation, or the attributes or actions of another person (e.g., Grover, 1997; Fulk & Mani, 1986; Neary Dunleavy, Chory, & Goodboy, 2010; O'Reilly & Roberts, 1974).

Deception at the Corporate or Macro-Level

As of February 2016, North American stock exchanges were valued at over $28 billion and represented 40.6% of the world total (Desjardins, 2016). According to a 2016 study conducted by the Association of Certified Fraud Examiners (ACFE), the "typical organization loses 5% of annual revenues to fraud" (Association of Certified Fraud Examiners, 2016, p. 4). Further, in their analysis of over 2400 fraud cases, the more individuals involved in the fraud, the more the money it cost the organization. Also, the higher the level of the conspirators in the fraud, the result was a higher level of loss (Association of Certified Fraud Examiners, 2016). Due to the impact of corporate-level deception, the following organizational variables will be reviewed: strategic ambiguity, corporate fraud, and deceptive advertising.

Strategic Ambiguity

First, it is important to dispel the belief that all deception is bad. As posited by Eisenberg (1984), "people choose communication strategies to accomplish multiple goals" (p. 238). He advocates that we move away from "an overly ideological adherence to clarity toward a more contingent, strategic orientation" (p. 239). In his article on strategic ambiguity, Eisenberg (1984) poses that strategic ambiguity is the purposeful use of ambiguous messages in organizations to accomplish goals. The goals of strategic ambiguity in organizations are that it "(1) promotes unified diversity, (2) facilitates organizational change, and (3) amplifies existing source attributions and preserves privileged positions" (Eisenberg, 1984, p. 227).

Strategic ambiguity promotes unified diversity by helping solve the paradox in organizations of promoting consistency and consensus yet encouraging creativity and independence of thought. Strategic ambiguity brilliantly allows for multiple voices to be heard but for organization members to feel as if they are still speaking the same language (Carmon, 2013).

Mission and vision statements can be used to examine strategic ambiguity. In Carmon's (2013) research on mission statements, for example, she found that of the 20 family business mission statements she analyzed, the consistent message was an openness to all employees, even those who were not family members. Strategic ambiguity in her research served to "create a sense of cohesion" for all employees in the organization (Carmon, 2013, p. 92).

The second goal of strategic ambiguity was to "facilitate change" (Eisenberg, 1984, p. 232). Organizations become successful because of their history but at the same time need to be open to changes in markets, consumer needs, and context or environment changes. An example comes from Ford Motor Company, which is considered by many to be a "family" business (Carmon, 2013) but has grown into an international corporation.

In 2017, the Ford Motor Company mission statement was: "*One Team. One Plan. One Goal*" and the vision statement was: "*people working together as a lean, global enterprise for automotive leadership*" (Thompson, 2017, emphasis in original, para 1; para 5). In March 2018, however, the Ford Motor Company webpage has a statement that appears to be a new mission statement, "At Ford, we go further to make our cars better, our employees happier and our planet a better place to be. Learn more about the work that makes Ford a company that we're proud to be a part of" (https://corporate.ford.com/company.html, 2018, para 1). The updated vision statement reads as, "People working together as a lean, global enterprise to make people's lives better through automotive and mobility leadership" (https://corporate.ford.com/company.html, 2018, para 2). These statements are strategically ambiguous as they emphasize unity and cohesion in the organization yet allow for independence and creativity of thought through their "One Plan," "make our cars better," and "automotive and mobility leadership." Ford employees know the goal is to change but they are encouraged to be creative in their contributions. The continued focus on a "lean" organization may be referring to an organization with fewer employees or one which embraces sustainability and more wisely uses resources so that the company can also make "our planet a better place to be." The mission and vision statements of the Ford Motor Company are inspiring as intended but they are ambiguous enough to warrant multiple possible meanings.

The final goal of strategic ambiguity is that it can be used to strengthen existing source attributions and preserve privileged positions (Eisenberg, 1984, p. 234). The key mechanism for supporting privileged positions is through the deniability of strategically ambiguous statements. If an

organizational member claims to have been misled by an ambiguous statement or metaphor, like "family," it is his/her *interpretation* of the statement or metaphor. Any interpretation can be plausibly denied. In the following examples of corporate fraud, although the fraudulent conspirators often employed complete distortions in their accounting deceptions, deniability of wrongdoing occurred, even if their efforts were clearly deceptive. Both of the following situations involved high-level executives who did not willingly communicate their actions and involved others through monetary promises.

Accounting or Corporate Fraud

Accounting or corporate-level fraud has been defined by the American Institute of Certified Public Accountants (AICPA) as "an intentional act that misstates the financial statements of an entity" (Jessup, 2013, p. 6). To better understand how this type of fraud occurs, we will look at it from the perspective of two well-known cases, Enron and the Madoff Ponzi scheme.

Enron

Enron was an organization which benefited from the deregulation of the energy markets, particularly the deregulation of the sale of natural gas (Lowry & Blinebry, 2014). The company grew exponentially and was considered by *Forbes* magazine to be the most innovative company in the US for 6 years (Stein, 2000). The company earned this reputation, in part, because of the creation of its online stock trading site—EnronOnline ("Enron Scandal," 2016).

Enron, like most organizations in the 1990s, was under great pressure to demonstrate consistent and competitive increases in their quarterly earnings (Berenson, 2003). At its most lucrative point, Enron shares sold for $90 a share (Edwards & Nadler, 2014). Even though the company appeared to be one of the best investments in the US, in 2001 Enron filed for Chapter 11 bankruptcy. This has been considered to be the biggest bankruptcy in US history and over 20,000 jobs were at stake because of it (Edwards & Nadler, 2014). By 2013, 5 of the allegedly involved executives were convicted and sent to prison (Windsor, 2013). Arthur Anderson, a respected accounting firm which served as auditors for Enron, took a major hit on its reputation because of the shredding of important Enron accounting documents, was indicted but later charges of obstruction of justice were reversed (Windsor, 2013).

Two questions most often asked when trying to understand what happened at Enron are: "How did they manage to pull this off?" and "How did they get caught?" Enron was successful at duping the financial world because of the complicated way they reported their losses. Enron, according to Berenson (2003), falsified both profits and losses. The company made many poor investments, and instead of reporting losses, the company "pretended to sell

the projects to partnerships of outside investors" (Berenson, 2003, p. 198), but the partnerships were Enron backed. Enron executives also employed a "gimmick called 'mark-to-market' accounting," which "enabled Enron to estimate what its future profits would be every time they signed a deal or made a trade – and then book those profits right away" (Berenson, 2003, p. 198). They counted potential future profits as current profits. All the fancy bookwork needed to perpetuate the fraud was accomplished through complicated reports investors could not understand. For example, when Kenneth Lay, Enron's then CEO, sold $70 million in stock, that should have been an indication of trouble at Enron. However, co-conspirators filed an obscure form called "Form 5" (Berenson, 2003, p. 199) to "report" the sale of the stock. Few investors understood or had heard of the form so it raised no concerns among investors at the time (Berenson, 2003).

The company also maintained a "corporate culture of deceit" (Lowery & Blinebry, 2014, p. 333). Executives, particularly the CEO Kenneth Lay, encouraged employees through bonuses while at the same time overlooking the breaking of rules and policies regarding financial operations (Lowery & Blinebry, 2014). The norm was to increase profits at any cost (Lowery & Blinebry, 2014). Enron's corporate culture was based on self-interest and greed.

As noted in an article by Park, Levine, McCornack, Morrison, and Ferrara (2002), most lies are detected not through the confrontation of those who lied but are revealed by third parties and physical evidence of the lie. Regarding Enron, individuals both at the auditing firm, Arthur Anderson (Bersenson, 2003) and Enron reported concerns about Enron's accounting. In particular, Sharon Watkins, who was an Enron Vice President at the time, wrote a memo to CEO Kenneth Lay regarding accounting concerns ("Enron Fast Facts," 2017). Journalists also called into question the ability of Enron to continue posting profits when other companies were struggling in a declining market (Lowery & Blinebry, 2014). With the announcement of a $638 million loss in October 2001 and an announcement by the SEC that Enron was being investigated, the fraud began to be revealed (Lowery & Blinebry, 2014). It was whistle-blowers and physical evidence that laid bare the fraudulent accounting processes of Enron.

Ponzi Scheme by Bernard Madoff
The term "Ponzi scheme" comes from Charles Ponzi who in the early 1900s promised US investors a 50% return on their investment in a stamp scam (Gossett, 2013). The promised return was not possible so Ponzi paid investors their return by bringing in new investors (Gossett, 2013). A Ponzi scheme occurs when investors are lured into investing based on these promises of high returns. The trick is that there is no true "investment" but money from new investors is used to pay "profits" to earlier investors (Gossett, 2013).

Bernard Madoff is an excellent example of corporate fraud and deception with his Ponzi scheme. He was believed to have been able to con thousands of investors, operating his Ponzi scheme for over 20 years (Gibson, 2014; Kirtzman, 2010). The Ponzi scheme Madoff ran resulted in stealing $17.5 billion in investments (Ross, Schwartz, & Christie, 2016) from investors who expected $65 billion in return for those investments (Yang, 2014).

Although Madoff claimed full responsibility for the scheme, he could not carry out the elaborate hoax without help. His sons were reportedly not involved in his scheme, but they were employed by his legitimate business (Gibson, 2014). In the end, Madoff was sentenced to 150 years in prison; his son, Mark, committed suicide two years after Madoff was arrested and his other son died of cancer in 2014; 14 co-conspirators pleaded guilty, 7 of whom cooperated with prosecutors and did not receive prison sentences, and 7 others received sentences ranging from 6 months to 10 years ("Irwin Lipkin…," 2015); and a little over $11 billion has been recovered (Ross et al., 2016).

As with Enron, much of the popular press articles and research analyses about Madoff ask the questions, "How did he do it?" and "How was he caught?" Gibson (2014) offers one of the most comprehensive reviews of the literature and the process by which Madoff and conspirators consistently fooled investors and put off the US Securities and Exchange Commission (SEC) investigators. Even with a whistle-blower, Harry Markopolos, another investment manager, the SEC persisted in its confidence in Madoff as one who had a "reputation as a successful businessman and philanthropist, which made him easy to trust and hard to challenge" (Gibson, 2014, p. 32). Madoff's reputation was one of the reasons posited by Gibson (2014) as to how the scheme was successful for so long. Other reasons Gibson (2014) puts forth include: Madoff had been successful as a legitimate fund manager before and during the time he ran the Ponzi scheme; he controlled information well, even going so far as having a separate, secret office where the scheme was managed and he gained promises from investors that they would not disclose their investments with Madoff; he was calm and seemingly open with SEC agents and journalists and he provided false portfolios as proof of his success; his clients believed he could bring in the returns he promised; and other Wall Street fund managers were skeptical but chose to not challenge Madoff (Gibson, 2014). Madoff, like the executives at Enron, also paid his conspirators well for their cooperation and perhaps, silence (Gibson, 2014).

According to research cited earlier by Park et al. (2002) and in the case of Enron, it was third parties and physical evidence which revealed Madoff's Ponzi scheme. Whistle-blowers such as Markopolos and journalists called the accounting practices of Madoff into question (Gibson, 2014). It was also the physical evidence which began to be exposed when the US market crashed in 2008, and investors started asking for their investments and returns

(Yang, 2014). Investors expected a return of $65 billion from Madoff, but he only had between $200 and $300 million he could pay back (Yang, 2014).

Advertising and Deception

Deception in advertising may not have the extreme impact as large-scale corporate fraud, as in the two previously discussed cases, yet deceptive advertising may mislead customers into paying more or buying products they do not need. The methods of distortion in advertising vary, but two of the most common types include puffery and covert, or subliminal marketing.

Examples of puffery can be found in nearly every advertising slogan. Puffery is "an ambiguous statement that includes a positive proclamation regarding a product" (Hubbell, 2014, p. 12). A current example of this is Papa John's Pizza slogan of "Better Ingredients, Better Pizza" ("Our Pizza," 2018, para 1). Puffery claims like "the best" or "better" are considered by the US Federal Trade Commission (FTC), as obvious exaggerations which reasonable customers would not believe and therefore not make buying purchases because of (Hubbell, 2014; Lee, 2014; Toncar, 2012); therefore, the FTC which is charged with monitoring deception in advertising, does not consider puffery to be deceptive (Hubbell, 2014; Toncar, 2012). Researchers, however, have found that puffery can influence buying decisions such that moderate to relatively high levels of puffery when coupled with a product trial are successful in influencing consumers' intent to buy the product (Hubbell, 2014). What is most ironic is that researchers have found effects of puffery with intent to buy, even with fictitious descriptions of items. For example, Carpenter, Glazer, and Nakamoto (1994) found that when research participants were presented with a completely irrelevant product description, an "alpine class down fill" for a down-filled jacket, the customers preferred this "alpine class down fill" jacket over identical jackets with accurate descriptions (p. 342). What was even more surprising in Carpenter et al.'s study was that even when participants were informed that the information was irrelevant, they still preferred the "alpine class" jacket to all others in the study. Therefore, although research continues to demonstrate that puffery influences buying intent, the FTC does not recognize it as deceptive (Toncar, 2012).

In contrast to overt claims as in puffery, the FTC considers some covert marketing as deceptive. Subliminal advertising, in particular, has been well researched (Hubbell, 2014). A subliminal advertisement is one where "consumers receive stimuli that predisposes them to rate a product more favorably, but they are not able to identify why" (Hubbell, 2014, p. 12). Colors, for example, have found to influence 60–80% of buying decisions (Markowitz, 2010). Blue, for example, is a color associated with cleanliness and is often used on soap or cleaning products, whereas red can get a product noticed but

is not a soothing color so may not be a good choice with products like contact solution (Markowitz, 2010).

The problem with deception in advertising, however, is in the determination of when an advertisement crosses the line into being duplicitous and determining if the deception was a purposeful act. For a message to be considered deceptive by the FTC, the advertisement must include false information that is crucial to convincing consumers to purchase the advertised brand or product. Dannon is a company which was ordered to pay customers $45 million through the result of a class action lawsuit. Dannon made claims "touting Activia and DanActive yogurt products as 'clinically' and 'scientifically' proven to regulate digestion and boost immune systems" (McMullen, 2010, para 1). These claims had no basis in fact and Dannon had to pull the advertising and settle the damages found in the lawsuit. Even though customers can file lawsuits for false claims in advertisements, the FTC is the major organization by which advertising campaigns are monitored. The FTC has also filed many lawsuits, even against deceptive online marketing practices. In 2012, the FTC requested and received "a judgment of more than $163 million on the final defendant in the 'Winfixer' scareware case where the defendants tricked consumers into thinking their computers were infected with malicious software in order to sell them software to 'fix' their nonexistent problem" (FTC, 2012, para 3).

How to Stop Deception in Organizations

Deception costs organizations money, can encourage consumers to buy products they may not want or need, and can result in people losing their jobs and/or life savings. This section will explore ways to encourage more honesty in organizations.

Code of Ethics

A code of ethics is often touted as the way to reduce deception in organizations. Publicizing values and expectations with regard to honesty and integrity are done by nearly every organization through the use of mission and/or vision statements.

However, some question the usefulness of a code of ethics, particularly when managers and/or executives in an organization act in ways that do not meet those codes of ethics (Pettit, Vaught, & Pulley, 1990). In the Enron case, for example, although there was reason to question Enron's success, the corporate culture was one where organizational members benefited financially from saying nothing (Lowery & Blinebry, 2014).

Also, organizations often create their code of ethics as a response to lawsuits or to try to avoid them. An organization may also be court mandated to establish and train employees on a code of ethics. The problem here is

the codes of ethics are established based on legal, although important criteria, instead of focusing on encouraging ethical behavior. These are called "compliance-based" codes of ethics (Paine, 1994). The Sarbanes-Oxley Act (SOX) became law in 2002 and represents a compliance-based code of ethics (Windsor, 2013). SOX was enacted as a response to the Enron disaster and other fraudulent organizational crises. SOX established more rules and regulations with regard to the financial management and auditing of organizations and they also created the Public Company Accounting Oversight Board (PCAOB) which established even more oversight of organizations (Windsor, 2013). In organizations, a compliance-based code of ethics may come from a professional standard, such as the SOX requirements, or may be based on previous lawsuits. If harassment has been an issue in an organization, that organization may include a no-tolerance policy of harassment in the next iteration of an organizational code of ethics.

In contrast, an "integrity-based approach" (Paine, 1994, p. 106) is a code of ethics that outlines expectations of positive behaviors (e.g., treat each other with respect), while including legal and ethical expectations (Paine, 1994). An example of the development of an integrity-based approach in response to customer concerns was discussed by Paine (1994). Paine uses the example of Sears, Roebuck & Company and complaints they had in 1992 with regard to their auto service division. Employees in that division were incentivized to sell parts and services. This had the unintended consequence of an increase in unneeded work on vehicles and customer complaints. According to Paine (1994), Sears changed their practice of incentivizing such sales in order to encourage more honest treatment of customers; they adapted behavioral expectations for employees from pushing sales to treating customers with integrity and honesty.

Integrity-based codes of ethics are constructed by first laying out the compliance-based codes or expectations. The legal demands and expectations of organizational members must be included in a code of ethics (Paine, 1994) as well as the values of the organizational members. Integrity-based codes can be developed by focusing on "the core values of integrity that reflect basic social obligations, such as respect for the rights of others, honesty, fair dealing, and obedience to law" and by "emphasize(ing) aspirations – values that are ethically desirable but not necessarily morally obligatory – such as good service to customers" (Paine, 1994, p. 112).

Outside Regulators of Businesses

Establishing a code of ethics is one way to encourage honesty and integrity in organizations. Another is to hold accountable those who actively deceive, commit fraud, or enact other deceptive behaviors. Regulatory agencies such as the SEC and FTC are charged with monitoring the behaviors of organizations; however, counting on these organizations to closely monitor every

organization is unrealistic. They often become involved when there is cause or complaints from competitors, consumers, or investors, or after a deceptive event, like what happened at Enron, has occurred.

Other potential checks and balances of organizations can be through those to whom information is disclosed including investors or outside directors. More oversight from outside directors can lead to more honest communication about negative organizational issues in annual reports. For example, in a study by Abrahamson and Park (1994), they conducted a content analysis of more than 1000 letters from corporate presidents in annual reports. They discovered that "outside directors, large institutional investors, and accountants limit such concealment but small institutional investors and outside directors who are shareholders prompt it" (Abrahamson & Park, 1994, p. 1302). Basically, Abrahamson and Parks (1994) found that more oversight by outside investors and directors led to organizational presidents revealing more negative information about an organization than organizational presidents with fewer outside investors and directors with whom they report.

Whistle-Blowers

Often, even with outsider oversight, only those within the organization know what really happens, making whistle-blowers crucial to outsiders becoming aware of problems. Whistle-blowing occurs when a previous or current organizational member reports wrongdoing in a company to someone inside or outside the organization who has the power or ability to address the wrongdoing (Fredin, 2011). Whistle-blowers have long been considered crucial to the detection and reporting of fraud or deceptive behaviors in organizations (Fredin, 2011), so much so that there is a "Welcome" to whistle-blowers on the SEC website ("Office of the Whistleblower," 2018). Whistle-blowers in the US are also protected by the Occupational Safety and Health Administration ("Your Rights as a Whistleblower," 2018), and federal employees are protected under the Whistleblower Protection Act of 1989 and Whistleblower Protection Enhancement Act ("Information on Whistleblower Protection Act and Whistleblower Protection Enhancement Act," 2017). However, even with protection, organizational members still weigh potential consequences before making a decision regarding reporting deception or other organizational wrongdoings (Fredin, 2011).

Deception at the Individual, Interpersonal, and Micro-Level

As discussed in the previous section, deceptive acts are achieved by the actions of one or more individuals. What makes this section different than the above section is the focus on relational variables, particularly those between a manager and subordinate that encourage or discourage deceptive messages and/

or behaviors. This section will focus on the research on topics such as motives for lying including variables such as trust, credibility, and power, as well as how employees lie.

Motives for Deception at Interpersonal or Individual Levels of an Organization

Deception in the workplace is commonplace, with up to 45% of employees admitting to having lied at work (Lindsey et al., 2011). Much of the research on deceptive communication among organizational members has applied an "interactionist perspective" (Grover, 1997, p. 72). Often, organizational members lie in response to a situation or context. Several situations or motives for lying previously explored have included role conflict (Grover, 1993, 1997), self-interest (Grover, 1997), rewarding deceptive behavior through incentives (Grover & Hui, 2005), pressure on employee to perform (Grover & Hui, 2005), employee moral development (Grover, 1993), organizational commitment (Grover, 1993), role stress (Fulk & Mani, 1986), and supervisory communication behavior and its impact on upward distortion (Fulk & Mani, 1986). Current research exploring organizational deception focuses on the supervisor-subordinate dyad in terms of the impact the interpersonal variables of trust, credibility, and power play into the relative dishonesty or honesty of organizational conversations. These interpersonal variables have been found to have a strong impact on deceptive behavior, particularly with upward communication (i.e., communication from subordinate to superior).

Trust

Trust is considered a "critical factor in affecting behaviors in organizations" (Nyhan, 2000, p. 88) because "no single variable influences interpersonal and group behavior as much as trust" (Sashittal, Berman, & Ilter, 1998, p. 163). When managers do not trust their associates, they must expend great amounts of energy and resources to monitor them (McAllister, 1995; Ruppel & Harrington, 2000), making trust one of the variables receiving much attention in organizational research (e.g., Neary Dunleavy et al., 2010).

An investigation of the behaviors that engender trust in organizations was undertaken by Whitener, Brodt, Korsgaard, and Werner (1998). Whitener et al. (1998) made use of agency and social exchange theories to develop a behavioral understanding of what engenders or creates trust. Using these theories, Whitener et al. (1998) posited that trust is based on the observation of another person's behavior(s). They argue that if a person acts in a trustworthy manner, then he or she can be trusted.

Whitener et al. (1998) used theory to arrive at five behaviors which induce trust in organizational colleagues. These behaviors were placed on

continuums, making the behaviors dimensional; individuals may demonstrate more or less of the behaviors. The behavioral dimensions included behavioral consistency (e.g., is the superior's behavior predictable?), behavioral integrity (including the superior communicating honestly), sharing and delegation of control (i.e., subordinates have some control over their lives/jobs), communication (i.e., open to feedback and open communication), and demonstration of concern (i.e., the superior would not harm the subordinate and has shown some concern for him or her) (Whitener et al., 1998).

Using the work of Whitener et al. (1998), a Managerial Trustworthy Behaviors Scale (Hubbell & Chory-Assad, 2005) was developed and used to determine how trust and perceptions of organizational justice (discussed in next section) influence antisocial behaviors, including a propensity to lie to a superior. A subordinate's level of trust in a supervisor was found to be an indicator of antisocial behaviors including a subordinate's propensity to be truthful with the supervisor (Chory & Hubbell, 2008; Hubbell & Chory-Assad, 2005).

The research on trust and its impact on relationships within organizational contexts has since been extended to co-worker dyads (Neary Dunleavy et al., 2010). Interestingly, in the co-worker dyadic research, findings indicate that how a person lies influences perceptions of that person. For example, in a hypothetical scenario study, participants were asked to evaluate a hypothetical co-worker saying one of many messages—some deceptive, one honest. The hypothetical co-worker who gave an honest message was considered more trustworthy by the research participants (Hubbell, Chory-Assad, & Medved, 2005). Researchers have also linked trust to other variables such as organizational justice.

Impact of Justice on Deception

Perceptions of injustice in the workplace have been found to bring about antisocial behaviors (e.g., Hubbell & Chory-Assad, 2005; Greenberg & Alge, 1998). Predictions regarding these effects have been based on equity theory (Adams, 1965; Walster, Walster, & Berscheid, 1978). Research on justice in organizations has demonstrated that perceptions of justice, or fairness, have been inversely related to behaviors which would be enacted to take revenge on the organization and/or employees in the organization (Colquitt, Conlon, Wesson, Porter, & Ng, 2001; Skarlicki & Folger, 1997).

Justice in organizations has been broken down into three types: distributive, procedural, and interactional justice. Distributive justice is a perception which results from the relative fairness of outcomes which result from an experience like a performance appraisal (Byrne & Cropanzano, 2001; Chory & Hubbell, 2008; Homans, 1961). When an individual evaluates distributive justice, he or she is comparing the outcome that he or she received to that of what another individual (e.g., a co-worker) received

with equity being the desired goal (Adams, 1965; Chory & Hubbell, 2008; Cropanzano & Greenberg, 1997). *Procedural justice*, in contrast, represents fairness perceptions of the procedures and processes that led to an outcome (i.e., performance appraisal and/or raise) (Byrne & Cropanzano, 2001). The fairness of process or procedures is evaluated "based on their consistency of application, prevailing ethical standards, the degree of their bias, accuracy, correctability, and the extent to which they represent all people concerned" (Chory & Hubbell, 2008, p. 359). Finally, *interactional justice* perceptions are related to actual treatment of an individual during a communication event. Respect, politeness, and maintaining another person's dignity are key to perceptions of interactional justice (Bies & Moag, 1986; Chory & Hubbell, 2008; Cohen-Charash & Spector, 2001).

In a study investigating perceptions of justice during performance evaluations, Chory and Hubbell (2008) found that when employees perceived low distributive and interactional justice, they were more likely to lie to superiors in future communication. When procedural justice was low, trust in a supervisor could mediate these perceptions of justice, thus reducing the propensity of lying as a result of low perceived procedural justice. Trust and justice have been found to be important when evaluating the likelihood of organizational members to deceive (Chory & Hubbell, 2008). Another key variable when considering organizational deceptive behaviors is power.

Power and Its Impact on Deceptive Behavior

Power is considered to be "one's ability to influence others" (Dunbar et al., 2014, p. 41). Power was broken down into power "bases" by French and Raven (1959) which include *reward power* (i.e., you can receive a reward by cooperating with the person with power), *coercive power* (i.e., the powerful person can punish you for a lack of cooperation), *legitimate power* (i.e., power is given because of an individual's position in a hierarchy), *referent power* (i.e., power is given because the person is someone you admire and whom you want to like you), and *expert power* (i.e., power earned through knowledge and/or abilities). Two power bases, *informational power* and *credibility*, have been added to the original five (Lindsey et al., 2011). Another form of power explored in organizational communication research is interactional dominance (Burgoon & Dunbar, 2000), which occurs when one individual attempts to control the behavior of another person (Dunbar et al., 2014).

The perceived power of a supervisor influences a subordinate's tendency to withhold relevant information from a supervisor or perform "gate-keeping" (Fulk & Mani, 1986, p. 484). The perception of how powerful a superior is influences what information he or she receives (Fulk & Mani, 1986; Read, 1962). Subordinates are also likely to communicate positive information as opposed to unfavorable information to their superiors when they believe that that superior has an upward influence in their organization

(Fulk & Mani, 1986; O'Reilly & Roberts, 1974). The importance of power in organizational deception has long been supported, yet has not generated extensive research (Lindsey et al., 2011). Therefore, four of the most influential articles linking power and deception in organizational contexts are discussed next.

First, the interactional aspect of deception and perceptions of power among co-workers were researched by Neary Dunleavy et al. (2010), Kelly (2015), and Dunbar et al. (2014). Neary Dunleavy et al. (2010) looked at the influence of honest or various deceptive messages on perceptions of co-workers' power, character, and trustworthiness. Neary Dunleavy et al. (2010) found the highest, or strongest, perceptions of competency, power, and trust for co-workers who communicated using truthful messages. However, when co-workers used either "withholding" or white lies as messages, as opposed to complete lies, they were perceived as still having relatively high levels of competence, character, expert power, and referent power. Similar findings were obtained with these types of messages by Hubbell and Medved (1999) and Kelley (2015). In an extension of Neary Dunleavy et al.'s work, Kelley (2015) found that managers who were perceived as being more deceptive were viewed as having less referent power and more coercive power.

Research by Dunbar et al. (2014) also explored perceptions of power and deception but their focus was on self-perceptions. Dunbar et al. (2014) designed a dyadic laboratory scenario where research participants were asked to make "mock hiring decisions" (p. 852) and participants were put into conditions where they were given hypothetical positions of reward power, equal power, or no power. Participants were also placed in deceptive or honest conditions. Findings indicated that regardless of power manipulation, those in the deceptive conditions perceived themselves as more powerful than their interactional partners. Participants who were instructed to be truthful reported perceiving themselves as having less power in the interaction. The deceptive communicators were also often perceived by their conversational partners as enacting more interactional dominance. Dunbar et al. (2014) came to the conclusion that participants in the deceptive condition, regardless of power manipulations, were more likely to perceive themselves as powerful and thus, were more successful in leading the hiring decisions made by the dyads.

Instead of co-worker relationships and deception, the final influential study of interactional deception and power to be examined here looked at the superior–subordinate relationship and deception. Ninety-six participants in Lindsey et al.'s (2011) study shared lies they had told either to their superiors, or as superiors they told to their subordinates. Although both groups reported that lies were not acceptable in an organization, both subordinates and superiors reported it was more acceptable for superiors, or those with more power, to lie, than those with less power (i.e., subordinates). As expected, participants also reported that in order to effectively deceive,

those with higher power were more likely to utilize their credibility and legitimate power, or authority. Those with less power relied on story-telling, often creating long and elaborate narratives and they were more likely to monitor their nonverbal behaviors than those with more power.

The act of lying in an organization is powerful (Carter, 2016). Power, justice, and trust are three of the most explored variables with regard to interpersonal or interactional deception in the workplace. Naturally, these variables are also closely related and may interact with one other and other dimensions, potentially creating a context where co-workers, superiors, and subordinates will be more likely to lie.

How to Reduce Interpersonal Organizational Lies

The four suggestions made in this section are based on previous research. They focus on proactive means of encouraging honesty in the workplace.

The first suggestion is to develop and follow ethical standards. One way to accomplish this is through an organizational code of ethics, one which employs both integrity- and compliance-based principles (Paine, 1994). Superiors or executives in organizations then have to follow that organization's code of ethics as the actions of superiors influence the actions of their subordinates. A superior who acts in an unethical manner and/or violated an organization's code of ethics communicates to subordinates that unethical behavior or violations of the organizational code of ethics are acceptable (Pettit et al., 1990).

The second suggestion stems from an understanding that the development of a code of ethics is not enough. Employees are more likely to lie when they endure role conflicts (Fulk & Mani, 1986; Grover, 1993), are pressured to perform beyond what is reasonably possible (Grover & Hui, 2005), and/or rewarded for achieving impossible or unattainable goals (Grover & Hui, 2005). With regard to role conflict, although not extensively studied in the current literature in organizational deception, when two superiors ask a subordinate to accomplish two conflicting assignments, the subordinate may feel compelled to pick one task from one supervisor and lie to the other supervisor about the other task (Grover, 1993). For example, if one supervisor asks an employee to move money from account A to B, and a different supervisor tells the same employee to never move the money from account A to B, the employee has to decide which action he will or will not take and then may feel compelled to lie to the superior he disobeyed.

Also, when individuals feel pressured to perform at a level that is not physically possible and are rewarded for such performance, they may feel compelled to lie (Grover & Hui, 2005). Therefore, close attention to the establishment of appropriate goal-setting, communication among managers pertaining to realistic performance, and achievable awards are important for keeping employees honest.

The third suggestion is to use theory and research on trust and justice to develop training programs and potential reward systems. For example, measures like the Managerial Trustworthy Scale (Hubbell & Chory-Assad, 2005) can be applied in any organizational context and easily adapted to assess the perception of superiors' or subordinates' enactment of trustworthy behaviors. After an organizational assessment of trustworthy behaviors, an organization could determine which behaviors are lacking among subordinates or superiors. Such information could then be used to develop training or incentive programs focused on encouraging behaviors, like integrity, that engender trust in others.

With regard to perceptions of justice, contexts such as the performance appraisal can be evaluated for employees' perceptions of fair and equitable treatment, using relevant instruments (see Chory-Assad, 2002; Chory & Hubbell, 2008; Hubbell & Chory-Assad, 2005). By evaluating perceptions of the differing types of justice, organizations may be able to improve procedures such as performance appraisals.

The fourth and final suggestion is to create and implement a system for reporting unethical or deceptive behavior. For example, one method employed by various organizations is called "EthicsPoint" ("NAVEX Global's EthicsPoint, n.d., para 1). EthicsPoint systems are usually operated by consultants who are not organizational members. Employees at any level within an organization can report either anonymously or confidentially to an EthicsPoint online system; note that the reporter's name is withheld from those whom they are reporting, but they may receive a report back on the result of their complaint. Upper-level management of an organization is informed of the complaint and has the responsibility to decide how to handle the situation or complaint.

FUTURE DIRECTIONS IN RESEARCH ON ORGANIZATIONAL DECEPTION

Much of the research in deception in organizations has focused at the interpersonal or interactional level. However, research on deception at the top of organizations, or the macro-level, has occurred primarily after major events such as the Enron or Madoff Ponzi scheme debacles. Research on what leads to more honest financial or other possible negative disclosures to shareholders is needed (e.g., Abrahamson & Park, 1994). Further, more examinations on the use of strategic ambiguity to achieve organizational goals may help in understanding the beneficial yet possibly deleterious impacts of strategic ambiguity (e.g., Carmon, 2013). More research is also needed to improve our understanding on how information is distorted and what compels organizations to be truthful. Continued research in organizational deception can offer support to those striving for the reduction in deceptive and fraudulent behaviors in organizations.

These directions lead to a second area for further exploration, which concerns the actual communication of deceptive messages. Several studies have used McCornack's (1992) information manipulation theory to explore perceptions of differing deceptive messages and honest messages (Hubbell & Medved, 1999; Hubbell et al., 2005; Kelley, 2015; Neary Dunleavy et al., 2010). The hypothetical scenarios used in these studies to examine perceptions of deceptive and honest messages could be expanded to look at more contexts or scenarios. It may be that in some contexts employees may view a complete lie as more acceptable than a white lie or even the truth. Current research has not yet found such scenarios but if found, they can help us better understand and work toward prevention of situations where employees feel compelled to lie.

Finally, integrity-based codes of ethics or conduct give organizations a way to focus on pro-social behaviors as opposed to antisocial behaviors (Paine, 1994). Instead of emphasizing problematic or antisocial behavior, an integrity-based code of ethics, combined with a compliance-based code of ethics, provides organizations with a more positive and proactive means of communicating values. Further, research linking the implementation of such codes of ethics with pro-social changes in behavior and the reduction in problematic behaviors could demonstrate the effectiveness of integrity-based codes of ethics.

In the end, we know that organizational members lie and sometimes those lies are strategically designed to help (Eisenberg, 1984), and sometimes those lies are used to conceal and ultimately hurt others. Through discussion of deception at two levels of an organization and in advertising, I hoped to have provided a more thorough understanding of the causes and consequences of organizational deceptive behavior. We may not be able to stop deceptive behavior in organizations but we may be able to encourage more truthful communication and control or minimize deceitful communication. When tragedies like fraud or even daily deceptive communication occur, it is usually the people at the bottom of the organization who are hurt the most. Through a discussion on deception in organizations and suggestions for future research, it is my hope that we can work toward reducing tragedies like Enron and Madoff's Ponzi scheme and improve the lives of those who suffer in deceitful organizational cultures.

References

Abrahamson, E., & Park, C. (1994). Concealment of negative organizational outcomes: An agency theory perspective. *The Academy of Management Journal, 37*(5), 1302–1334. http://www.jstor.org/stable/256674.

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 267–299). New York: Academic Press.

Association of Certified Fraud Examiners. (2016). *2016 Global fraud study*. Retrieved from https://www.acfe.com/rttn2016/docs/2016-report-to-the-nations.pdf.

Berenson, A. (2003). *The number: How the drive for quarterly earnings corrupted Wall Street Corporate America*. New York: Random House.

Bies, R. J., & Moag, J. S. (1986). Interactional justice: Communication criteria of fairness. *Research on Negotiation in Organizations, 1,* 43–55.

Burgoon, J. K., & Dunbar, N. E. (2000). An interactionist perspective on dominance-submission: Interpersonal dominance as a dynamic, situationally contingent social skill. *Communication Monographs, 67,* 96–121. https://doi.org/10.1080/03637750009376497.

Byrne, Z. S., & Cropanzano, R. (2001). The history of organizational justice: The founders speak. In R. Cropanzano (Ed.), *Justice in the workplace: From theory to practice* (Vol. 2, pp. 3–26). Mahwah, NJ: Lawrence Erlbaum Associates.

Carmon, A. F. (2013). Is it necessary to be clear? An examination of strategic ambiguity in family business mission statements. *Qualitative Research Reports in Communication, 14*(1), 87–96. https://doi.org/10.1080/17459435.2013.835346.

Carpenter, G., Glazer, R., & Nakamoto, K. (1994). Meaningful brands from meaningless differentiation: The dependence on irrelevant attributes. *Journal of Marketing Research, 31*(3), 339–350. https://doi.org/10.2307/3152221.

Carter, M. (2016). Deceit and dishonesty as practice: The comfort of lying. *Nursing Philosophy, 17,* 202–210. https://doi.org/10.1111/nup.12129.

Chory-Assad, R. M. (2002). Classroom justice: Perceptions of fairness as a predictor of student motivation, learning, and aggression. *Communication Quarterly, 50,* 58–77. https://doi.org/10.1080/01463370209385646.

Chory, R. M., & Hubbell, A. P. (2008). Organizational justice and managerial trust as predictors of antisocial employee responses. *Communication Quarterly, 56*(4), 357–375. https://doi.org/10.1080/01463370802448121.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes, 86,* 278–321. https://doi.org/10.1006/obhd.2001.2958.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86,* 424–445. https://doi.org/10.1037//0021-9010.86.3.425.

Cropanzano, R., & Greenberg, J. (1997). Progress in organizational justice: Tunneling through the maze. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (Vol. 12, pp. 317–372). London: Wiley.

Desjardins, J. (2016, February 17). *All of the world's stock exchanges by size.* Retrieved from http://www.visualcapitalist.com/all-of-the-worlds-stock-exchanges-by-size/.

Dunbar, N. E., Jensen, M. L., Bessarabova, E., Burgoon, J. K., Bernard, D. R., Harrison, K. J., … Eckstein, J. M. (2014). Empowered by persuasive deception: The effects of power and deception on dominance, credibility, and decision making. *Communication Research*, *41*(6), 852–876. https://doi.org/10.1177/0093650212447099.

Edwards, T., & Nadler, J. T. (2014). Business. In T. R. Levine (Ed.), *Encyclopedia of deception*. Thousand Oaks, CA: Sage.

Eisenberg, E. M. (1984). Ambiguity as strategic in organizational communication. *Communication Monographs, 51,* 227–242. https://doi.org/10.1080/03637758409390197.

Enron scandal: The fall of a Wall Street darling. (2016, December 2). Retrieved from http://www.investopedia.com/updates/enron-scandal-summary/.

Federal Trade Commission (FTC). (2012). *Challenging deceptive advertising and marketing.* Retrieved from https://www.ftc.gov/reports/annual-report-standard/ftc-2013/challenging-deceptive-advertising-and-marketing.

Ford Motor Company. (2018, March 4). Retrieved from https://corporate.ford.com/company.html.

Fredin, A. J. (2011). The effects of anticipated regret on the whistleblowing decision. *Ethics and Behavior, 21*(5), 404–427. Retrieved from http://dx.doi.org/10.1080/10508422.2011.604296.

French, J. R. P., & Raven, B. (1959). The basis of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 529–569). Ann Arbor, MI: University of Michigan Press.

Fulk, J., & Mani, S. (1986). Distortion of communication in hierarchical relationships. *Communication Yearbook, 9*, 483–510. Retrieved from http://dx.doi.org/10.1080/23808985.1986.11678621.

Gibson, D. R. (2014). Enduring illusions: The social organization of secrecy and deception. *Sociological Theory, 32*(4), 283–306. Retrieved from https://doi.org/10.1177/0735275114558631.

Gossett, J. (2013). Ponzi schemes. In L. W. Salinger (Ed.), *Encyclopedia of white-collar and corporate crime* (Vol. 2, pp. 731–733). Thousand Oaks, CA: Sage.

Greenberg, J., & Alge, B. J. (1998). Aggressive reactions to workplace injustice. In R. W. Griffin, A. O'Leary-Kelly, & J. M. Collins (Eds.), *Dysfunctional behavior in organizations: Violent and deviant behavior* (Vol. 23, Part A, pp. 83–117), *Monographs in organizational behavior and relations.* Stamford, CT: JAI Press.

Grover, S. L. (1993). Why professionals lie: The impact of professional role conflict on reporting accuracy. *Organizational Behavior and Human Decision Processes, 55,* 251–272.

Grover, S. L. (1997). Lying in organizations: Theory, research and future directions. In R. Giacalone & J. Greenberg (Eds.), *Antisocial behavior in organizations* (pp. 68–84). Thousand Oaks, CA: Sage.

Grover, S. L., & Hui, C. (2005). How job pressures and extrinsic rewards affect lying behavior. *International Journal of Conflict Management, 16*(3), 287–300. https://doi.org/10.1108/eb022933.

Homans, G. C. (1961). *Social behavior.* New York: Harcourt, Brace & World.

Hubbell, A. (2014). Advertising, consumer products. In T. R. Levine (Ed.), *Encyclopedia of deception.* Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781483306902.n4.

Hubbell, A. P., & Chory-Assad, R. M. (2005). Motivating factors: Perceptions of justice and their relationship with managerial and organizational trust. *Communication Studies, 56*(1), 47–70. http://dx.doi.org/10.1080/0008957042000332241.

Hubbell, A. P., & Medved, C. (1999). *The integration of three perspectives of organizational deception: A typology of deceptive messages.* Paper Presented at the National Communication Association Annual Convention, Chicago.

Hubbell, A. P., Chory-Assad, R. M., & Medved, C. E. (2005). A new approach to the study of deception in organizations. *North American Journal of Psychology, 7,* 171–180.

Information on whistleblower protection act and whistleblower protection enhancement act (2017, July 14). Retrieved from https://www.sec.gov/eeoinfo/whistleblowers.htm.

Irwin Lipkin gets 6 months in prison, ending Bernie Madoff case. (2015, August 5). Retrieved from http://www.nbcnews.com/business/business-news/irwin-lipkin-last-defendant-madoff-case-faces-sentencing-n403971.

Jessup, C. M. (2013). Accounting fraud. In L. M. Salinger (Ed.), *Encyclopedia of white-collar and corporate crime* (Vol. 1, pp. 6–9). Thousand Oaks, CA: Sage.

Kelley, P. J. (2015). *Deception among organizational leaders: Impacts on employee perceptions of supervisor credibility, power, and trust* (Unpublished Master's thesis), University of Tennessee, Knoxville. Retrieved from http://trace.tennessee.edu/utk_gradthes/3378.

Kirtzman, A. (2010). *Betrayal: The life and lies of Bernie Madoff*. New York: Harper.

Lee, S. (2014). When do consumers believe puffery claims? The moderating role of brand familiarity and repetition. *Journal of Promotion Management, 20*(2), 219–239. https://doi.org/10.1080/10496491.2014.885481.

Lindsey, L., Dunbar, N., & Russell, J. (2011). Risky business or managed event? Perceptions of power and deception in the workplace. *Journal of Organizational Culture, Communications and Conflict, 15*(1), 55–79.

Lowery, M., & Blinebry, A. (2014). Enron. In T. R. Levine (Ed.), *Encyclopedia of deception*. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781483306902.n128.

Markowitz, E. (2010). *How to choose the right colors for your brand*. Retrieved from https://www.inc.com/guides/2010/10/how-to-choose-the-right-colors-for-your-brand.html.

McAllister, D. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *The Academy of Management Journal, 38*(1), 24–59. http://www.jstor.org/stable/256727.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs, 59*, 1–16. https://doi.org/10.1002/9781118540190.wbeic072.

McMullen, T. (2010, February 26). *Dannon to pay $45 M to settle yogurt lawsuit*. Retrieved from http://abcnews.go.com/Business/dannon-settles-lawsuit/story?id=9950269.

NAVEX Global's EthicsPoint. (n.d.). Retrieved from https://secure.ethicspoint.com/domain/en/default_reporter.asp.

Neary Dunleavy, K., Chory, R. M., & Goodboy, A. K. (2010). Responses to deception in the workplace: Perceptions of credibility, power, and trustworthiness. *Communication Studies, 61*(2), 239–255. http://dx.doi.org/10.1080/10510971003603879.

Nyhan, R. C. (2000). Changing the paradigm: Trust and its role in public sector organizations. *American Review of Public Administration, 30*, 87–109.

Office of the Whistleblower. (2018, March 4). Retrieved from https://www.sec.gov/whistleblower.

O'Reilly, C. A., & Roberts, K. H. (1974). Information filtration in organizations: Three experiments. *Organizational Behavior and Human Performance, 11*, 253–265.

Our Pizza. (2018, March 4). Retrieved from https://www.papajohns.com/company/how-we-make-better-pizza.html.

Paine, L. S. (1994). Managing for organizational integrity. *Harvard Business Review, 72*(2), 106–117.

Park, H. S., Levine, T., McCornack, S., Morrison, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs, 69*(2), 144–157. https://doi.org/10.1080/714041710.

Pettit, J. D., Vaught, B., & Pulley, K. J. (1990). The role of communication in organizations: Ethical considerations. *International Journal of Business Communication, 27*(3), 233–249. https://doi.org/10.1177/002194369002700302.

Read, W. H. (1962). Upward communication in industrial hierarchies. *Human Relations, 15*, 3–15.

Ross, B., Schwartz, R., & Christie, M. (2016, February 1). *Recovered Madoff money now over $11 billion; More hidden away?* Retrieved from http://abcnews.go.com/US/recovered-madoff-money-now-11-billion-hidden/story?id=36529402.

Ruppel, C. P., & Harrington, S. J. (2000). The relationship of communication, ethical work climate, and trust to commitment and innovation. *Journal of Business Ethics, 25*(4), 313–328. https://link.springer.com/article/10.1023/A:1006290432594.

Sashittal, H. C., Berman, J., & Ilter, S. (1998). Impact of trust on performance evaluations. *Mid-Atlantic Journal of Business, 34*, 163–185.

Skarlicki, D. P., & Folger, R. (1997). Retaliation in the workplace: The roles of distributive, procedural and interactional justice. *Journal of Applied Psychology, 82*, 416–425.

Stein, N. (2000, October 2). *The world's most admired companies*. Retrieved from http://archive.fortune.com/magazines/fortune/fortune_archive/2000/10/02/288448/index.htm.

The fall of Enron. (2017, July 15). Retrieved from http://www.npr.org/news/specials/enron/.

Thompson, A. (2017, February 5). *Ford Motor Company's vision statement, mission statement*. Retrieved from http://panmore.com/ford-motor-company-vision-statement-mission-statement.

Toncar, M. (2012). A study of visual puffery in fragrance advertising: Is the message sent stronger than the actual scent? *European Journal of Marketing, 46*(1/2), 52–72. https://doi.org/10.1108/03090561211189239.

Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research*. Boston: Allyn & Bacon.

Whitener, E. M., Brodt, S. E., Korsgaard, M. A., & Werner, J. M. (1998). Managers as initiators of trust: An exchange relationship framework for understanding managerial trustworthy behavior. *Academy of Management Review, 23*, 513–530. http://www.jstor.org/stable/259292.

Windsor, D. (2013). Enron corp. In L. W. Salinger (Ed.), *Encyclopedia of white-collar and corporate crime* (pp. 303–304). Thousand Oaks, CA: Sage.

Yang, S. (2014, July 1). *5 years ago Bernie Madoff was sentenced to 150 years in prison—Here's how his scheme worked*. Retrieved from http://www.businessinsider.com/how-bernie-madoffs-ponzi-scheme-worked-2014–7.

Your Rights as a Whistleblower. (2018, March 4). Retrieved from https://www.osha.gov/OshDoc/data_General_Facts/whistleblower_rights.pdf.

Physician Deception and Telling the Truth About Medical "Bad News": History, Ethical Perspectives, and Cultural Issues

H. Russell Searight and Taylor Meredith

Determining if an act is deceptive requires assessment of both motive and context (Stearns, 2014). In healthcare settings, patients may report non-existent symptoms with the goal of monetary gain through a disability settlement or to obtain more favorable treatment in the prison system (Fitzgerald & Danner, 2014). Pharmaceutical companies have been accused of withholding information about potential conflicts of interest when they underwrite medical education programs that are, in reality, venues for promoting the company's new medication (Hubbell, 2014). In medical research, such as in double-blinded clinical trials, participants are randomly assigned to an active pharmaceutical agent or placebo; neither the patient nor the clinical investigator knows the identity of those receiving either "treatment." However, in trials in which it is necessary to disclose potential medication side effects, the clinical researcher's description of these adverse events may artificially increase their probability in the placebo group through the power of suggestion (Blease, 2015). This chapter focuses on another issue in medical care in which deception and deliberate withholding of information occur,

H. Russell Searight (✉)
Department of Psychology, Lake Superior State University,
Sault Sainte Marie, MI, USA
e-mail: hsearight@lssu.edu

T. Meredith
Lake Superior State University, Sault Sainte Marie, MI, USA

specifically with regard to informing patients of a serious, often terminal, illness.

Deliberately withholding information about serious and terminal illness from patients had, until approximately 50 years ago, been a common, accepted practice in medicine in the US (Everett et al., 2011; Sisk, Frankel, Kodish, & Isaacson, 2016). Frequently, physicians either did not inform patients of serious medical conditions such as cancer or used euphemisms such as "mass" or "growth" to indicate cancerous tumors. In addition to intentionally withholding information, physicians were often overtly optimistic about prognoses even when they were aware of patients' imminent death (Lamont & Christakis, 1999), practices that continue to be evident in recent times (Smith, Dow, Khatcheressian, & Lyckholm, 2010). Historically, the shift to "truth telling" in US medicine was a gradual process influenced by legal rulings, changes in the physician–patient relationship, and greater attention to established ethical theory in medical education. From the perspective of medical ethics, non-disclosure was consistent with principles of non-maleficence (i.e., "do no harm") and a paternalistic beneficence (i.e., what is best for the patient). However, the most widely accepted ethical model in contemporary medicine emphasizes patient autonomy, which is predicated on informed consent.

While the emphasis on autonomy and a collaborative, rather than authority-based, physician–patient relationship is considered the norm among Americans of White Northern European background, cross-cultural research suggests that these values are not universally held (Larkin & Searight, 2014). Research has indicated greater variation in preferences for patient-centered versus family-centered decision-making among several ethnic communities in the US (Rising, 2017; Searight & Gafford, 2005a). A common theme among those who do not subscribe to a norm of complete informed consent and patient-centered decision-making is that it is emotionally and possibly physically harmful to an individual already struggling with illness (Rising, 2017; Searight & Gafford, 2005a). In addition, disclosure of medical news to a loved one may be viewed as disrespectful to parents and grandparents (Garcia-Prerto, 2016).

DEFINING AND CONTEXTUALIZING MEDICAL "BAD NEWS"

While it may be surprising, physicians have often seen it as appropriate, even therapeutic, to deliberately withhold information from patients about serious and life-threatening medical conditions. Diagnoses of cancer, Alzheimer's disease, and kidney failure, as well as results of biopsies have either been explained in very oblique ways (e.g., "you have a small mass on your liver") or with a distorted picture of the prognosis (e.g., "patients have been known to live years with a pancreas that is not functioning normally"). Additionally, for over half of the previous century, it was common for physicians not to inform patients

of serious illness. If patients asked directly about the seriousness of their condition, a good physician would "…skillfully mix…falsehood and truth" (Collins, 1927, p. 320). Another physician strategy was to overtly ignore the reality of the patient's condition:

> Doctors' ready retreat behind silence—apparent to patients by doctors' demeanor when they keep most of their thoughts to themselves, deprive patients of vital information, or pat patients on the back and assure them that everything will be all right. (Katz, 2002, pp. 209–210)

In the US, Canada, and Northern Europe, the practice of providing deceptive descriptions of serious medical conditions, or simply avoiding discussion of the topic with patients, has changed significantly in the past 50 years. However, as law and ethics progressively influenced medical practice in recent decades, there are still concerns about whether these disclosures of "bad news" are in patients' best interests. Medical "bad news," a phrase that became popular in the medical literature, was defined in 1984 by Buckman as "…any information likely to alter drastically a patient's view of his or her future" (p. 1597).

Beginning in the 1980s, the issue of cultural competence in healthcare was becoming a standard part of medical school education. Physicians, nurses, and other healthcare professionals soon found that many patients and their families, particularly those from non-Western cultures, did not view patient autonomy as empowering, but as isolating (Rising, 2017). To make a truly autonomous decision about one's healthcare required complete information about one's condition, prognosis, and treatment options. However, to many immigrant communities, the US standard of receiving complete diagnostic information was viewed, at minimum, as harsh. Candib (2002) describes an exchange between a physician and a Russian immigrant patient as the patient was told of his cancer diagnosis:

> Initially [the patient] listened carefully with a rather stoic demeanor. I was certain that my approach would again be the right one. However, as the news sank in, he became rather agitated and appeared quite angry. This prompted me to have the interpreter explore his feelings: [Medical doctor] (MD): You look very upset and angry. Can you tell me what you are thinking? [Patient] (Pt): I am very angry that you have given me such terrible news. Where I come from, doctors give bad news to the family. MD: So, you are angry because I told you about the likelihood of having cancer. Pt (appearing angrier): Yes, and I don't want to talk with you about it anymore. You are telling me that I have a terrible disease without chance of cure. How am I to get better without having hope? Please just decide what you need to do and talk with my family and Dr. Danilova (the patient's regular physician). They will tell you what to do. MD: (rather remorseful and shocked): I am sorry if I have upset you. I believe that it is important for people to be in charge of their healthcare and only wanted

to provide you with that opportunity. The patient acknowledged the explanation with a dismissive "yes" and made no further eye contact. (Candib, 2002, pp. 215–216)

HISTORY OF DISCLOSING MEDICAL "BAD NEWS" TO PATIENTS

The ancient Greeks, including Hippocrates (460–370 BC) as well as Renaissance-era physicians, did not have a clear understanding of disease and prognosis, and only had a limited number of treatments. While the historical record is somewhat ambiguous, it appears that Hippocrates placed himself on the side of non-disclosure with his suggestion that knowledge of serious illness could further harm the patient while indicating that relevant information should be disclosed to a third party (Krisman-Scott, 2000). The Ancient Greeks are typically seen as an enlightened exception in their approach to medicine (Martin, 2015). The Middle Ages (400–1400 AD) were characterized by religious explanations of illness as the product of sin. With the Enlightenment during the 1600s and 1700s, reasons and science were applied to medicine; however, concepts such as germ theory were not recognized until the late 1800s (Martin, 2015). In the 1800s, the norm appeared to be against disclosure. In 1803, in a work considered the first comprehensive presentation of medical ethics, Percival (2014) indicated that physicians were not required to disclose negative medical information if it could potentially be harmful to the patient. Percival's work became the foundation for the American Medical Association's first ethics code (Beauchamp & Childress, 2013). However, a minority of physicians expressed concern that deceiving patients could result in public mistrust of the medical profession (Hooker, 1849; Krisman-Scott, 2000). Nevertheless, during the past century, legal and clinical issues, as well as questions about directly informing patients about serious illness, treatments, and accompanying risks, have become much more prominent. To illustrate how the norm of non-disclosure shifted to a policy of sharing "bad news" with patients, a brief history, covering the last 100 years, is presented next.

1920–1950

In the early twentieth century, several physicians wrote articles for the public on the question of deceiving patients for therapeutic reasons (Eisenberg, 1986; Sisk et al., 2016) with titles such as "A Patient Wants to Know" (Yost, 1936). In a 1927 article in the popular magazine, *Harpers*, Collins argues against truth telling based on the social and emotional harm that the information may have on seriously ill patients. He contends that because of their learned compassion, truth telling is an impossible task for the physician: "To tell the truth is often to perpetuate a cruelty of which many are incapable" (p. 320).

In his article, Collins (1927) confesses to an error in judgment in which he disclosed a serious medical condition and immediately saw the harm it caused the patient. In this situation, an attorney with pain in the shoulder and arm verbally minimized the significance of his symptoms. The patient was found to have serious kidney disease, and Collins's recommendations to the patient included retiring from his job, reducing physical activity, and significant dietary change. Collins had not yet disclosed the significance of the condition; however, the patient pressed him, and Collins explained the seriousness of the disease. Collins describes the patient's response: "the light of life begins to flicker from the fear that my words engendered, and within two months it sputtered and died out. He was the last person in the world to whom the truth should have been told. Had I lied to him, and then intrigued with his family and friends he might be alive today" (p. 322).

Collins's regret about being honest is palpable—he describes his truthfulness with the patient as akin to medical malpractice. With few exceptions, Collins believes that those who are indeed terminally ill are those who are least likely to directly ask about their diagnosis and prognosis. He indicates that patients who "honestly and courageously" want to know the truth are extremely rare and that these patients are guided by the desire to "face the wages of sin while there is still time" (p. 320).

Collins provides advice to physicians of the era: "...cultivate lying as a fine art" (p. 322). By the 1940s, some physicians were cautiously and tentatively disclosing serious diagnoses, and always with the objective of trying to maintain hope. Providing patients with this information was not necessarily seen as helpful or beneficial; the treatment of the issue typically centered around whether these disclosures could be harmful. In a 1935 *New England Journal of Medicine* article, Henderson discussed the role of social sciences in medicine, describing patient–physician interaction as a dynamic system. While one could take issue with his belief that science had not progressed beyond the work of Machiavelli in understanding "...the influence of sentiments upon the actions of men..." (Henderson, 1935, p. 819), Henderson did appreciate the ambiguity in situations involving physician disclosure of medical bad news. Acknowledging that both disclosing and withholding diagnostic information could be harmful to patients, Henderson emphasized the role of the physician's professional judgment about a patient's response to bad news.

The 1950s

Henderson's interest in physician–patient interaction was the exception until the latter part of the twentieth century. It was not until the 1970s, with the patient rights movement and the growing consumer orientation among recipients of medical services, that physician paternalism began to be challenged (Rothman, 1991; Tomes, 2016). As such, it is not surprising that while the

medical profession viewed disclosure as harmful to patients (Eisenberg, 1986; Kaufman, 1953), the general population's opinion had rarely, if ever, been systematically examined (Eisenberg, 1986; Fox & Swazey, 2008). A survey of 477 patients published in *Time* magazine found that 96.5% wanted to be told if they had cancer and 88.5% wanted close relatives to be informed (Bowen, 1954).

Physicians, however, if aware of these data, appeared not to trust these findings. Echoing Collins's conclusions from the 1920s, it was one thing to ask people in the abstract if they would like honest information about a medical condition (Bowen, 1954); however, it was an entirely different circumstance to be asked these questions in the reality of current or imminent pain and suffering.

During the same decade, two surveys of physicians' perspectives on informing patients seemed to contradict each other. A survey of over 5000 physicians residing in Philadelphia found that 70% of physicians never informed patients of a cancer diagnosis, while 30% indicated they "always" or "usually" disclosed the diagnosis (Fitts & Ravdin, 1953, p. 904). However, in a smaller sample of 108 Wisconsin surgeons, 71 (66%) reported that they always or usually informed patients of cancer, 20 (19%) indicated that they sometimes informed patients, and 17 (16%) indicated that they "seldom or never" disclosed a cancer diagnosis (Maloney, 1954, p. 548). However, the surgeons pointed out that informing patients of their condition was often required since surgery was being recommended. Further, the availability of new treatments, often with significant side effects, and required patient adherence over extended time periods, made it difficult for physicians to allay patients' fears with platitudes such as "It's nothing serious."

While the requirement of patient consent to medical procedures had been legally established in the US in 1914 in *Schloendorff v. Society of New York Hospital*, the specific information necessary for patients to make these decisions was not clear. In 1957, in the case of *Salgo v. Leland Stanford University Board of Trustees*, the court ruled that the essential risks of a medical procedure must be disclosed to patients prior to surgery. This ruling, based on a patient who indicated that he had not been told of potential risks and whose surgery left him paraplegic, is often credited with the establishing of the phrase, "informed consent" (Krisman-Scott, 2000; Rothman, 1991).

The 1960s

In the 1960s, the issue of informed consent in biomedical research received public attention, while survey findings indicated that clinicians were still reluctant to disclose medical bad news. Oken's 1961 survey of 219 physicians suggested that non-disclosure continued to be standard practice. Of those surveyed, 90% indicated having concealed a cancer diagnosis at some point in their careers.

Challenges to practices in biomedical research also had implications for a standard of informed consent in clinical practice. In 1966, Beecher, an anesthesiologist, published what is now considered a landmark study in biomedical research ethics in the *New England Journal of Medicine*. Beecher (1966) concluded that much medical research was being conducted without explicit informed consent from participants. In his exposé, Beecher noted that only two of the 50 articles that he reviewed, which were published in prestigious medical journals, explicitly mentioned consent (Beecher, 1966; Rothman, 1991). In the 22 cases that Beecher described in detail, there were some studies in which deception appeared to be fairly explicit:

> Example 17 from Beecher's 1966 study: [Under medical supervision] "…Live cancer cells were injected into 22 human subjects as part of a study of immunity to cancer. According to a recent review, the subjects (hospitalized patients) were "merely told they would be receiving 'some cells' …The word cancer was entirely omitted". (Beecher, 1966; cited in Rothman, 1991, p. 74)

However, despite analogizing contemporary biomedical research practices to the Nazi experiments described in the Nuremberg Tribunal and legal rulings (Beauchamp & Childress, 2009; Fox & Swazey, 2008; Rothman, 1991), deception and withholding diagnostic information was still common practice in clinical research. Even 20 years later, more than 50% of the physicians surveyed indicated that it was appropriate not to disclose to a patient that they had been enrolled in a clinical cancer trial and were receiving a placebo (Faden, Beauchamp, & King, 1986).

While Beecher's exposé focused on research, the concept of informed consent received additional legal support in a 1960 clinical case (*Natanson v. Kline*) involving a patient who had not been adequately informed of the risks of receiving radiation treatment for cancer (Rothman, 1991). Ms. Natanson, after undergoing a radical mastectomy, received cobalt radiation therapy. The radiation treatment resulted in significant injury to the chest, including skin, cartilage, and bone. A key element of the plaintiff's case was that Dr. Kline, the radiologist, had not adequately warned her of the possible negative outcomes of radiation therapy. The case was eventually heard by the Kansas Supreme Court, which recognized that there were circumstances in which a physician's professional judgment not to provide detailed information about a recommended treatment could be justified. However, in Ms. Natanson's case, there was no acceptable reason for withholding information about treatment risks (Plante, 1967).

The 1970s

Kübler-Ross's (1926–2004) text *On Death and Dying* (1969) became well known in the 1970s. In her interviews with more than 400 patients over a

four-year period, she found that patients, when informed of a terminal illness, experience a series of reactions or stages. In her original formulation, Kübler-Ross (1969) presented these responses to learning of a terminal diagnosis as a specific sequence: denial, anger bargaining, and depression acceptance. Kübler-Ross, Wessler, and Avioli (1972) recount how they had requested that physicians who had patients who were dying in a large hospital refer them to the study—after a week "there was not a single dying patient in that 600-bed hospital" (p. 12). Her observation illustrates the role of physicians' anxiety and helplessness in non-disclosure. However, Kübler-Ross noted that the discomfort adversely influenced physicians' communication, so that they came across like "stammering suitor[s]" (pp. 13–14).

Of the patients that she interviewed, approximately 40% indicated that they had never been told directly about their prognosis. Kübler-Ross concluded, however, that the majority of these patients knew that they were indeed dying—often observing family members' stilted conversations, friends' forced cheerfulness even though it was evident they had been crying, and the fact that once their diagnosis was known to be terminal, nurses required twice as long to respond to the call button (Kübler-Ross et al., 1972).

Kübler-Ross et al. (1972) advise when patients ask directly, they should never be told that they are dying. Instead, they advise that patients should be "spared the worst" with phrases such as "it looks pretty grim" or "it looks pretty bad" (Kübler-Ross et al., 1972, p. 175). They also advise that the clinician should wait for the patient's response to gauge how they are coping with the information. While Kübler-Ross's stages are not as universal and linear as originally described, her work was very important in bringing discussions of death into medical education.

By the end of the decade, there appeared to be evidence of significant change in physicians' views of disclosure. In a study using similar methodology and survey questions as Oken (1961), Novack et al. (1979) found astonishingly different responses. Specifically, 97% of physicians indicated that they typically informed patients of a cancer diagnosis. However, as Krisman-Scott (2000) notes, the response rates may have had a significant role in these studies. Oken's (1961) study had a 90% response rate; Novack et al.'s (1979) rate was approximately 25%. However, anecdotal reports suggest that the norm of deception and stonewalling silence was still pervasive in patient care (Mullan, 2016).

The 1980s

During the 1980s, physicians favoring disclosure were making significant headway. However, much of the impetus was not based on ethics, but by law and institutional policy. The rise in malpractice litigation, which not only impacted physicians but also support personnel such as nurse anesthetists and hospitals, was probably a major contributor toward informed consent as a "medical Miranda warning" (Meisel & Kuczewski, 1996, p. 2523).

Descriptions of the recommended procedure, risks, benefits, and alternatives were presented verbally—often in a monotonic rush—and accompanied by a form, resembling a contract, for the patient to sign. This perfunctory approach often did not achieve the ethical objectives of informed consent since it often did not stimulate a dialogue between physician and patient about the patient's illness, prognosis, and risks and benefits of treatment options (Searight & Barbarash, 1994).

Informal observations and descriptions in the literature (Callahan, 1992) suggested that many physicians disclosed to their patients begrudgingly and with a tone of being coerced by legal advice or bureaucratic policy (Beauchamp, 2011; Searight & Barbarash, 1994; Searight & Miller, 1996). Thus, these disclosures tended to be technical and a medical procedure in and of itself. The ritual of obtaining the signature on the consent document became another medical task, which permitted emotional distancing by the physician (e.g., often called *consenting* the patient, as in "Can you consent the patient before we start the procedure?").

These legalisms also led physicians to engage in "truthful lying" (Epstein, Korones, & Quill, 2010; Meisel & Kuckewski, 1996). For example, a detailed risk-benefit analysis in a medical situation in which there may be ambiguity about diagnosis and few well-established evidence-based treatments, while consistent with fully informing the patient, may not be helpful—particularly if the patient was experiencing emotional distress (Politi, Lewis, & Frosch, 2012). As documenting patient informed consent became an established requirement during the 1980s, some physicians disclosed information to meet their legal requirements, but at the same time much of their verbiage was not comprehensible to the patient, and physicians were aware of these communicative issues (Meisel & Kuckewski, 1996). Language such as "neoplasm," "sarcomas," "elevated laboratory values," and "shadows on x-rays" are examples of medical terminology often used to describe details of a patient's condition. However, words such as these conveyed incomprehensible information to a reasonable adult without medical training (Weir, 1980). This procedural approach has been described as "truth dumping" (Callahan, 1992; Smith, 1988). This technical, detached approach to patient disclosure also became ritualized in the implementation of the Patient Self-Determination Act (PSDA) of 1990, which is discussed later.

Weir (1980), an ethicist on the faculty of the University of Iowa College of Medicine, writing to physicians, while recognizing that some physicians were reluctant to adopt full disclosure, described potential benefits. Weir (1980) recognized that from a legal standpoint, disclosure was now required but also believed in a continued role for physician discretion: "Telling the truth and being truthful with patients can be considered…A contextually limited privilege granted by physicians when they consider it appropriate and withheld when they judge it not in the best interests of a particular patient" (p. 208).

Weir (1980), also raised the issue of informed consent in the context of medical care. He argued that medical care was becoming more specialized,

consumer-oriented, and less personal. Further, he contended that physicians often did not know their patients well enough to know whether or not "benevolent deception" was appropriate (p. 211).

1990–2010

While disclosure of bad news to patients had become a standard by the 1980s, communication with patients continued to be nuanced. Gordon and Daugherty (2003) reported results with a focus group of oncologists and found that while they did disclose to patients that cancer was often incurable, they were very reluctant to convey a specific prognosis. These oncologists also reported their concern about the intense psychological damage that grim prognostic information would have for patients. While they wanted patients to have "realistic expectations" about treatment, they also experienced an obligation to maintain patient hope (Gordon & Daugherty, 2003).

With growing attention to research on physician–patient communication (Stewart, 1995), as well as medical education (Smith, 2002), the particularly difficult task of conveying "bad news" became a topic of interest to medical educators and clinicians. Quill and Townsend (1991) described specific objectives for the conversation: including reducing feelings of isolation for both patient and physician; arriving at a common perception of the problem; providing the information needed by the patient for imminent decisions; helping the patient develop a short-term plan; and assessing for suicide. Importantly, they pointed out that the physician should not draw conclusions prematurely about the meaning of the diagnostic and prognostic information for the patient but should make significant efforts to understand it from the patient's perspective (Quill & Townsend, 1991).

Quill and Townsend's (1991) work has been cited widely and adapted to various medical specialties. Consequently, delivery of medical bad news has become a standard component of medical education in doctor–patient communication (Rosenbaum, Ferguson, & Lobes, 2004). Delivering medical bad news has become so integral in medical education that trainees are taught an acronym to remember its key components. Rather than relying on observations of senior physicians during clinical education or personal trial and error, resident physicians and medical students have been taught an acronym when delivering bad news: SPIKES (Setting, Perception, Invitation, Knowledge, Empathize, Summary, and Strategy) (Baile et al., 2000). While this mnemonic device may seem mechanistic and cold, protocols are a common way of organizing information within medical culture.

Recent Re-Examinations of Non-Disclosure of Medical Bad News

By 2010, physicians who had been advocates and proponents of patient autonomy and "truth telling" such as Quill recognized that unmitigated

disclosure, despite being legally and ethically required, might not always enhance patient autonomy (Beauchamp & Childress, 2013; Politi, Lewis, & Frosch, 2012). Medical informed consent waivers might be useful in situations in which patients do not want to be informed (Beuachamp & Childress, 2009; Searight, 1992; Searight & Barbarash, 1994). The waiver, more commonly used in medical research, is a legally recognized document that involves two distinct waivers—waiving information surrounding one's medical condition and waiving decision-making about treatment (McKinney et al., 2015; Meisel, 1979).

While some patients are not cognitively able to understand these waivers, must receive urgent medical attention, or the patient explicitly indicates that they do not wish to be informed (Epstein et al., 2010), other situations are ambiguous and require clinical judgment to determine the appropriateness of disclosure. Epstein and colleagues observed that patients who frequently change the subject or are silent in response to the clinician's information may be communicating the desire to not be informed. Posing even greater ambiguity are "gray areas" in which a patient does exhibit intact cognition and has not clearly expressed a desire not to be informed of their diagnosis or treatment options (Searight & Gafford, 2005a). In these situations, physicians should reflect upon whether having this medical information would benefit the patient (Pellegrino & Thomism, 1993).

Too much information can also be indirectly harmful to patients. When patients are offered several treatment options, a recitation of all conceivable adverse events or side effects can cognitively obscure the most relevant dimensions that need to be weighed in their decision-making (Epstein et al., 2010; Epstein & Peters, 2009; Truog et al., 2015). Whenever numerous medical tests are conducted, the likelihood of a positive, yet clinically insignificant result is high (Truog et al., 2015; Yang, Scarfe, & Angelopoulos, 2018). Epstein and colleagues (2010) present the example of a patient with a potentially life-threatening lung nodule who is also found to have a benign liver cyst. If presented with both findings, the patient may be unable to prioritize the significance of the two conditions. In fast-paced medical environments where time is limited and costly, not addressing a relatively insignificant medical issue allows the patient and physician to focus on decision-making about life-threatening conditions.

Epstein et al. (2010), who had previously written about the importance of presenting patients with "bad news" and the best methods for these disclosures, raise the possibility that indiscriminate sharing of medical information might reduce patient autonomy. Providing in-depth information about illness or treatment options may challenge patients' cognitive processing skills (Baile et al., 2000)—specifically the ability to prioritize information that is relevant to one's values and life goals. While not providing clear evidence, Epstein et al. (2010) suggest that patients who are persistently seeking more detailed information from a physician may be signaling problems with

trust. For example, there is often a justifiable level of mistrust of the White-dominated medical establishment that continues to be perceived as racist (Feagan & Bennefield, 2014; Washington, 2006). This distrust has a long history dating back to the early to mid-1800s when slaves were subjects of experimental surgeries conducted without anesthesia (Owens, 2018), as well as in the twentieth century when African-American men with syphilis were denied treatment as part of studies of the course of the disease (Jones, 1992; Reverby, 2009; Washington, 2006).

Finally, in making the judgment to withhold medical information, the ethical maxim of fidelity to patient well-being should be the overarching principle (Pellegrino & Thomasma, 1993). Ideally, a physician will appreciate their patients' values and be able to promote patient autonomy by including these dimensions when discussing diagnostic and treatment options (Epstein et al., 2010).

MEDICAL ETHICS AND THE QUESTION OF DISCLOSING BAD NEWS

The 1970s saw the beginning of contemporary medical ethics (Beauchamp & Childress, 1979, 2009). As medicine became more complex with interventions such as organ transplants, advanced technology including dialysis and mechanical ventilation, and hospital intensive care units, the moral dimensions of healthcare also became more apparent. The Hippocratic Oath, an early code of physician ethical conduct from ancient Greece (approximately 400 B.C), which has since undergone revisions, has traditionally been taken by new medical school graduates at commencement (Markel, 2004). However, by the late 1960s, the oath was no longer sufficient to guide the types of decisions that healthcare professionals had to make themselves and with their patients. Professional ethicists, often trained in philosophy, beginning in the late 1970s and early 1980s, became part of university medical school faculty (Fox & Swazey, 2008). Historians of medical ethics often emphasize different historical developments in their accounts. However, frequently cited developments include the ability to successfully transplant organs. The rise in transplant surgeries contributed to discussion about the definition of life since comatose patients' vital functions could be maintained with technology. Another precipitating factor was the initially limited access to kidney dialysis. The practice of heart and kidney transplants raised issues of possible conflicts of interest between donors, patients, and surgeons (Rothman, 1991). The moral dimensions of these clinical dilemmas lay outside of biomedical science (Fox & Swazey, 2008). As a result, other disciplines, such as philosophy, were brought into medicine.

However, for healthcare providers seeking a protocol or definitive guide to decision-making about what to tell patients, the guidance provided by professional ethicists was often relativistic and certainly not definitive. While older clinicians had likely become sensitized to a view of medical ethics as a unitary

reflection of the work of Hippocrates (Edelstein, 1967), professional ethicists, invoking philosophy, presented different models of ethical reasoning each with its own implications. The most commonly cited models and their implications for disclosure of "bad news" are briefly discussed below.

Kant's Deontology

Kant believed that there were moral absolutes, sometimes called categorical imperatives. These are universal laws that are essentially nonnegotiable. Hartmann (1932), an ethicist, presents the Kantian perspective of absolutes: "…truthfulness…admits of no expectation at all…no end can justify deliberate deception as a means" (p. 194). Kant recognizes, at some level, that there may be situations in which individuals experience a duty to lie to avert a harmful outcome. However, according to Kant, deception is, without qualification, morally wrong. Even if the physician does not tell the patient the entire truth about their condition because of concerns about the emotional harm it may have, one is still left with the guilt associated with lying. For Kant, there is no escaping this responsibility for violating truthfulness: "Whatever militates against frankness, lowers the dignity of man [sic]" (translation by Infeld, 1936, p. 154).

There are clinicians who still operate from moral absolutes—physicians who refuse to assist hastened death in a terminally ill, dying patient suffering in pain. Similarly, other physicians following moral absolutism refuse to perform abortions, while others reject prescribing contraception. When physicians operate in absolutist fashion, they behave according to the Kantian doctrine. From Kant's perspective, non-disclosure of a serious illness equates to deception (translation by Infeld, 1936). To Kant, the motive behind deception is irrelevant, even if it is benevolent and carried out with a duty to protect the patient from harm (Pence, 2014). If one's duty is to be honest, there are no compromises based upon situational factors, including the patient's emotional well-being (Johnston & Holt, 2006). Physicians who report always telling patients of their terminal illness, such as in the case example of the Russian patient described earlier (Candib, 2002), are behaving in accord with Kantian deontology.

Utilitarianism (Consequentialism)

Utilitarianism, heavily influenced by Jeremy Bentham (1748–1832) and John Stuart Mill (1806–1873), focuses on the outcome of deception or withholding information rather than the act itself (Pence, 2014). An action by a healthcare professional resulting in the greatest benefit to the greatest number of people is considered to be the ultimately correct action. From a utilitarian perspective, there are no acts that are considered universally right or wrong. One can only evaluate the morality of an act in terms of its consequences.

Utilitarian thinking is implicitly involved in any type of rationing system. For example, the availability of kidney dialysis in the 1960s led to one of the most controversial applications of utilitarianism (Jonsen, 2007). The number of individuals who would benefit from kidney dialysis was far greater than the availability of dialysis machines at Washington State's Swedish Hospital. A committee was appointed by the county medical society that included a cross section of citizens including a lawyer, minister, housewife, state government official, labor leader, and surgeon. The "God Committee," as it came to be known (Alexander, 1962), made a series of utilitarian judgments. After making initial utilitarian judgments based upon patients' age—specifically, children (which were not clearly defined) and persons over age 45—were screened out of the pool. Next, a list of characteristics including age, sex, marital status, net financial worth, psychological stability, educational background, occupation, and other criteria were used to determine who was most "worthy" of dialysis (Jonsen, 2007). The public description of the "God Committee" in *Life* magazine (Alexander, 1962) likely led to the program of federal support for dialysis and its greatly expanded availability.

A utilitarian rationale has also been invoked as a reason for withholding harmful information (Collins, 1927; Epstein et al., 2010). The potential harm of the emotional and perhaps physical reaction to this difficult information outweighs the value of honesty or truthfulness. Depending on the situation, utilitarianists would agree with Collins (1927) that cultivating the fine art of lying might indeed be one of the most morally useful assets of a physician. If, in the larger context of the patient's life, being spared knowledge that would be emotionally distressing reduces the aggregate harm experienced in the patient's lifetime, a utilitarian ethic of deception would be considered morally correct.

Historically, when the rationale for non-disclosure has been to prevent patients' psychological distress, the courts have been inconsistent in their views of the utilitarian or consequentialist perspectives (Johnston & Holt, 2006). Careful reading of legal decisions suggests an upper threshold of distress that may permit non-disclosure. A 1980 Supreme Court case in Canada, *Reibl v. Hughes*, serves as a particularly important example. In particular, a patient, while undergoing surgery, suffered a stroke that left him partially paralyzed. Reibl argued that he had not been given complete informed consent indicating that the physician had not clearly articulated the risks and benefits of having versus not having the surgery. The physician indicated that the patient had a life-threatening condition and may have not consented to the procedure if serious, yet low probability risks were described. The Canadian courts concluded that "...a particular patient may, because of emotional factors, be unable to cope with facts relevant to recommended surgery or treatment and the doctor may, in such a case, be justified in withholding or generalizing information as to which he would otherwise be required to be more specific" (Johnston & Holt, 2006, p. 147).

However, a review of legal rulings on informed consent (Walter, 1997) suggests that being emotionally upset is, alone, not enough to justify non-disclosure. In another ruling, not temporally distant from *Reibl v. Hughes*, the court, while recognizing emotional distress as a concern, expressed an almost-Kantian argument in holding the physician liable:

> It may well be that he [the doctor] considered the claimant over-anxious or over preoccupied with "horror stories" and the possibility of being crippled. In these circumstances I do not find it improbable that, in an attempt to reassure, he deflected her inquiries by answering them in [the] light-hearted term… However understandable such a response may have been in psychological terms, it was not an adequate response in legal terms. (*Chester v. Afshar*, 2005; cited in Johnston & Holt, 2006, p. 148)

While paternalistic physicians have made a case for utilitarianism, the rise of a consumer orientation among patients over the past 30 years (Tomes, 2016), coupled with an increasingly litigious culture, makes utilitarianism difficult to defend legally. Additionally, it is difficult to reconcile the emphasis on patient autonomy with the value of protecting patients by non-disclosure of medical bad news.

Virtue Ethics

Virtue ethics, exemplified by the Hippocratic Oath, shares with Kant's deontology a focus on the clinicians' acts rather than the impact on the patient. The virtue perspective is reflected in the list of the actions and motivations of a "good" person. In the case of healthcare, virtue ethicists, when confronted with a moral dilemma, ask "What would a good doctor do?" (MacKenzie, 2009, p. 196). The Hippocratic Oath is written as a narrative of desirable and moral behaviors that an ethical physician exhibits (e.g., "Into whatever houses I enter, I will go into them for the benefit of the sick and will abstain from every voluntary act of mischief and corruption; and further from the seduction of females or males, of freemen or slaves"). As noted earlier in the discussion of history, Hippocrates viewed the virtuous physician as one that would protect the patient from "needless" emotional distress.

In contemporary medical ethics, Pellegrino and colleagues (e.g., Pellegrino & Thomasma, 1987) are probably most closely tied to the Hippocratic tradition in that they emphasize the moral character traits of "good" physicians. These characteristics and accompanying duties include intellectual honesty, compassion, fortitude, temperance, integrity, and self-effacement, as well as benevolence and humility. A core virtue is *phronesis*, described best as prudent wisdom. By exercising this form of thoughtful reflection, the physician will be able to choose the moral course of action in the face of clinical dilemmas. These duties are not simply actions to be carried out, and virtue ethicists

would certainly disagree with the "procedural" approach to informed consent to satisfy legal obligations. Even though it is the "right thing to do," duties cannot be isolated from the physician's overall character. Virtue ethics, although not as untethered from the consequences of physicians' decisions on patients as deontology, is not necessarily utilitarian. Contemporary virtue ethicists have attempted to reconcile doing what is best for the patient with the promotion of patient autonomy. However, critics (e.g., Beauchamp & Childress, 2009) question whether the paternalism of non-disclosure to protect a patient from emotional harm can be congruent with the promotion of autonomy.

Pellegrino and Thomasma (1987) argue that beneficence, acting on the patient's behalf, typically involves enhancing patient autonomy. However, as Epstein and colleagues (2010) have noted from a clinical perspective, full disclosure of medical information may not necessarily enhance patient autonomy, especially for patients who characterologically have difficulty separating "big picture" issues from less significant minutiae. When confronted with a situation in which the physician is reluctant to disclose a particularly pessimistic prognosis, a useful exercise, consistent with a virtue perspective, is to consider what a respected colleague would do in the situation.

Principlism

Beauchamp and Childress (2009) reviewed existing philosophical approaches to medical dilemmas and distilled them into four key principles: autonomy, beneficence, non-maleficence, and justice. Principlism, sometimes referred to as "the four principles", is the dominant approach to medical ethics in the US (Page, 2012). The first three dimensions are most relevant in clinical situations; justice (i.e., treating others equally) is an important dimension in public health and distribution of healthcare resources. However, it is not a central feature in issues surrounding "truth telling" in physician–patient interactions.

Autonomy has been equated with self-determination; it reflects one's ability to control their future through decisions and corresponding action. Patient autonomy is violated or compromised when individuals do not have all relevant information about decisions impacting their lives. A major criticism of principlism that is evident in "truth telling" and other clinical dilemmas is that it does not specify priorities among the principles; all four dimensions are given equal weight.

However, many commentators have indicated that autonomy is "first among equals" (Searight, 2016). Beauchamp and Childress (2009), however, assert otherwise and indicate that principlism does not assume any priority or ranking of the principles. This assumption of equality among the principles, however, appears contradictory when viewed in the context of some of the

major issues in bioethics in the past 20 years. Issues such as physician-assisted suicide, informed consent, the right to die, advance directives, and patient competence for medical and financial decision-making, all center around patient autonomy.

When conducting medical ethics consultations in a general hospital, the first author (Searight) organized written findings in terms of the four principles. However, in several cases of seriously ill patients who could not communicate, there was a search for the patient's voice—often difficult to elucidate when they were comatose or experiencing delirium. "What would Mr. Smith want in this situation?" was the question frequently raised in the quest for the autonomous decision-maker, whose ability to express a choice was obscured by illness. Written advance directives were encouraged so that autonomous choice could be expressed when the individual was no longer able to do so.

Non-maleficence, doing no harm, is often invoked in the controversy surrounding physician-assisted death or suicide (PAS). Physicians opposing PAS view a patient's death as the "ultimate harm" and in a distinct moral category separate from patient distress associated with terminal illness. In principlism, the emotional and possibly indirect physical harm to a patient associated with receiving a diagnosis of serious illness is one of the major arguments for non-disclosure. The patient has already been harmed by an illness that is typically outside of their control. The principle of non-maleficence argues that it is wrong to add the harm of psychological distress to an already suffering patient. However, dating back to the nineteenth century, physician and ethicist Worthington Hooker argued that physicians who withhold the truth from patients perpetuate another type of harm (Beauchamp & Childress, 2009; Hooker, 1849). Specifically, he maintained that adopting a policy of deceiving patients with serious illness would erode public trust in the medical profession.

Beneficence, the moral obligation to benefit others, has often been fused with non-maleficence (i.e., to do good is not to harm) (Beauchamp & Childress, 2009). While Beauchamp and Childress note meaningful distinctions between the two principles, healthcare professionals view non-disclosure as benefiting patients by not harming them further. Even among healthcare professionals who believe in telling patients that they have a serious condition, an honest statement about a patient's prognosis is often that the outcome is unknown. A physician stating that a patient's prognosis is unknown, while honest, also conveys hope (Beauchamp & Childress, 2013). Further, the outcome of the course of the illness and response to treatment are often unpredictable. The reality is that physicians do not know the "whole truth" (Beauchamp & Childress, 2013). Physicians' knowledge limitations of patients' prognoses, as well as the ambiguity of clinical medicine, permit only an approximation of the objective truth (Beauchamp & Childress, 2009).

The Impact of Culture on Disclosures
of Medical Bad News

From an ethical perspective, cross-cultural research on communication about serious illness and treatment decisions suggests that the prevailing US norm of direct patient disclosure may not be sensitive to the values of many cultures (Searight & Gafford, 2005a). This cultural perspective, reflecting a communitarian ethical model, suggests that communitarian values may, in part, be a legacy of a shared history. For example, in the US, the Tuskegee syphilis study is well known in the African-American community. Beginning in the 1930s, a group of African-American males from rural Alabama were enrolled in a study conducted by the US Public Health Service (Reverby, 2009; Washington, 2006). The overall objective of the study was to understand the course of syphilis. However, this goal was not disclosed to the participants who believed that they were receiving treatment when they underwent regular blood draws and periodic spinal taps. Even when penicillin was established as an effective treatment for the condition in the 1940s, the men in the study were not informed of its availability. Many contemporary ethicists have argued that the legacy of Tuskegee includes a pervasive suspiciousness toward the medical establishment associated with this history of deception. African-American patients are less likely to be organ donors, less likely to participate in clinical trials, and have very low rates of completion of advanced directives (Searight & Gafford, 2005a). Even when all of the elements of informed consent are disclosed, there may well be lingering distrust (Boulware, Cooper, Ratner, LaVeist, & Powe, 2003).

From an ethical perspective, communitarians have been critical of individualism and an accompanying absence of duty or responsibility to others. Indeed, moral responsibilities to others are an infringement upon individual autonomy and freedom (Theobold & Wood, 2009). Rather than promoting the common good, the emphasis on autonomy reflects the increasingly influential norm of hyperindividualism (Theobold & Wood, 2009) in the US (Bellah, Madsen, Sullivan, Swidler, & Tipton, 2007). However, a communitarian perspective is seen in cultures in which it is common for the family, rather than the individual patient, to receive and act upon medical information (Rising, 2017).

Non-Disclosure Across Cultures: Research

In the mid-1990s, the *Journal of the American Medical Association* featured two studies—one qualitative (Carrese & Rhodes, 1995) and the other a large sample quantitative study (Blackhall, Murphy, Frank, Michel, & Azen, 1995)—investigating views of medical decision-making and advance directives among ethnic communities in the US. These studies reflected responses to the recently implemented PSDA, which required that all

institutions receiving Medicare or Medicaid funds provide patient education about advance directives. Advance directives are of two basic types—a living will that specifies specific levels of care that one desires if unable to express their wishes, and a durable power of attorney which names someone (e.g., typically a spouse, adult child, or other family member) to make decisions on the person's behalf. Legislative attention to advance directives was, at least, partially the result of several well-publicized cases of adults (Pence, 2017). One such case included Nancy Cruzan, who was comatose for multiple years, and whose family had made decisions regarding life support that were not followed by medical personnel (Pence, 2017). The key issue was the absence of a clear record indicating the patient's preference for continued life support in situations where recovery was highly unlikely.

In keeping with its status as a federal health agency, the Bureau of Indian Affairs implemented the PSDA in its medical facilities. Carrese and Rhodes (1995) noted that these discussions were not well received in some Native American communities. Death and poor medical outcomes were, by tradition, not openly discussed. Carrese and Rhodes (1995) described the response to the PSDA from Navajo elders: "You don't say those things. And you don't try to bestow that upon yourself…The object is to live if possible here on earth. Why try to shorten it by bestowing things upon yourself?" (Carrese & Rhodes, 1995, p. 828). In the Navajo community, it was customary to "talk around" these issues, often by speaking in generalities or sharing a narrative about someone else's experience when seriously ill. Not being told about the reality of one's medical condition, such as having cancer or a serious cardiovascular disease, was expected and not considered deceptive.

Soon after the initiation of the PSDA, Blackhall et al. (1995) published a large sample quantitative study on cross-cultural views of end of life decision-making. Surveying a pool of 800 individuals—with equal representation among Korean-, European-, Mexican-, and African-American participants—researchers inquired about participants' views of physician disclosure of a cancer diagnosis (Blackhall et al., 1995). While a clear majority of European- and African-American respondents indicated that patients should be informed, 47% of Korean- and 65% of Mexican-Americans agreed with informing patients of a cancer diagnosis. Additionally, 57% of Korean-Americans and 45% of Mexican-Americans who participated in the study believed that the family, rather than the patient, should make the decision about life support technology (Blackhall et al., 1995).

Several qualitative studies suggest that there are at least three explanations for limiting disclosure and for family-centered decision-making.

Talking About Death Brings It Closer: As the scenario described by the Navajo elder makes clear (Carrese & Rhodes, 1995), in developing both advanced directives and living wills, individuals are often asked to imagine being seriously ill and on life support. The patient may be queried about their preferences for continued treatment when chances of survival are poor,

unknown, or if they are unable to communicate a decision. In cultures such as the Navajo, in which language, thought, and action are inextricably linked (Carrese & Rhodes, 1995), end of life discussions are not hypothetical future possibilities but, instead, give terminal illness and death a reality and possibly make it a self-fulfilling prophecy. In Carrese and Rhodes's (1995) study, several of the members of the Navajo community declined to discuss advance care planning because of the associated belief that these discussions could be dangerous. Participants stressed the importance of thinking and speaking in the "beauty way." The view that verbalizing, even hypothetically, about illness, shapes reality and may make future events more likely is expressed by this Navajo woman's description of her father's experience:

> The surgeon told him that he may not wake up, that this is the risk of every surgery. For the surgeon, it was very routine, but the way that my Dad received it, it was almost like a death sentence, and he never consented to the surgery. (Carrese & Rhodes, 1995, p. 828)

The belief that being told about a possible negative medical outcome can be physically or psychologically harmful, while not as explicit as among the Navajo, has been found in other cultures such as recent Bosnian immigrants to the US (Searight & Gafford, 2005b). These respondents characterized advance directives as "playing with your destiny."

Filial Piety, Love, and Respect: In many Asian cultures, it is considered disrespectful to require an elder to grapple with a cancer diagnosis and make decisions about treatment (Searight & Gafford, 2005a). Rather than simply protecting an elderly family member from emotional distress, non-disclosure occurs in a web of mutual obligation and responsibility. The below exchange between a Korean-American woman and an interviewer provides a glimpse of the complexity when making end-of-life decisions for family members (see Frank et al., 1998). Life-sustaining treatment is not solely for the patient, but must also occur with great respect for others:

> *Interviewer*: If the patient were you then what do you expect your children would decide for you?
> *Interviewee*: If my children wanted to see me even one more day, then they might ask for the treatment. I am the one who is going to die; so, I don't control the situation.
> *Interviewer*: When you think about the situation right now, would you want the treatment for your life to be extended if you were not conscious and had almost no hope to live?
> *Interviewee*: I would rather pass away sooner if by having my life extended it caused pain.
> *Interviewer*: But for others, you would ask for the treatment to extend the life?
> *Interviewee*: In other cases, if the patient were either my child or my husband then I would request the treatment to see them even a little longer.

Interviewer: Isn't that contradictory?

Interviewee: Although it's a contradiction it's the right thing to do…Don't you think so? (Frank et al., 1998, p. 411)

This is an example of filial piety, or the combined respect and love for an older family member, which is another explanation for limiting disclosure and for family-centered decision-making. Deception, as an act of care toward a family member, is illustrated by a story that became popular in China after the news media regarded it as the epitome of romantic love (Nie, 2013). The story described a man who was married for 17 years and with one child was diagnosed with lung cancer (Zhao, 2015). His body did not appear to respond to treatment. His wife, concerned that he would feel hopeless and depressed about his condition, obtained his laboratory reports and made copies for her husband; however, on the copies, she deliberately changed the laboratory test numbers to indicate improvement rather than deterioration. Moreover, in the event that her husband asked the physician for test data, she maintained a set of "optimistic" results at the physician's office. Rather than being criticized for deceiving or manipulating her husband, the wife was widely praised for her "true love" and devotion (Nie, 2013).

Maintaining Patient Hope and Preventing Harm: In many cultural communities, disclosing bad news is not seen as promoting autonomy, but as increasing feelings of depression and hopelessness. One theme that arose among the study of Bosnian immigrants was the perception that for those who were already in pain, a physician who disclosed an emotionally painful truth was being cruel (Searight & Gafford, 2005b). An important part of the physician's role, according to these participants, was to maintain patients' hope. "[In Bosnia]…the doctors—always say there's hope; there's different techniques…Hope was always given to the patient" (Searight & Gafford, 2005b, p. 199). Even if some degree of deception was necessary, maintaining the patient's hope was far more important than the patient's informed consent.

Conclusion

Historically, deception and non-disclosure have been guided by concern for patient welfare. From Hippocrates to the modern era, non-maleficence and beneficence are the ethical principles upon which physicians' decisions to avoid further troubling patients with "bad news" are based. In the long-standing paternalistic practice of medicine, the physician's judgment about not telling a patient of a cancer diagnosis was acceptable and supported—at least within the medical community. However, the rise of patient rights, consumerism, and in many instances, loss of an enduring physician–patient relationship have all contributed to patients being more active participants in making their own healthcare decisions.

From an ethical perspective, the key conflict in disclosing or failing to disclose serious medical information to patients centers on autonomy versus beneficence. While recognizing the importance of patient autonomy, including treatments to restore patients' ability to make independent decisions, Thomasma and Pellegrino (1987) note that a key moral duty of physicians is to promote patient self-determination. Genuine respect for persons includes recognizing and supporting the decision to be uninformed and to delegate decision-making to a trusted physician and/or family member. Healthcare professionals who genuinely "demonstrate respect for persons" (Thomasma & Pellegrino, 1987, p. 45) will preface any disclosures with a respectful query about whether the information is desired and if the patient would like someone else to receive information and/or make medical decisions on their behalf. Patients are always reminded that if they would like their diagnostic and treatment information, the physician will readily provide these details (Searight & Gafford, 2005a). Allowing patients to choose whether to be informed respects their autonomy while not compromising the healthcare professional's integrity.

References

Alexander, S. (1962). They decide who lives, who dies. *Life, 53,* 102–125.

Baile, W. F., Buckman, R., Lenzi, R., Glober, G., Beale, E. A., & Kudelka, A. P. (2000). SPIKES—A six-step protocol for delivering bad news: Application to the patient with cancer. *Oncologist, 5,* 302–311.

Beauchamp, T. L. (2011). Informed consent: Its history, meaning, and present challenges. *Cambridge Quarterly of Healthcare Ethics, 20*(4), 515–523.

Beauchamp, T. L., & Childress, J. F. (1979). *Principles of biomedical ethics.* New York: Oxford University Press.

Beauchamp, T. L., & Childress, J. F. (2009). *Principles of biomedical ethics* (6th ed.). New York: Oxford University Press.

Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics* (7th ed.). New York: Oxford University Press.

Beecher, H. K. (1966). Ethics and clinical research. *New England Journal of Medicine, 274,* 1354–1360.

Bellah, R., Madsen, R., Sullivan, W., Swider, A., & Tipton, S. (2007). *Habits of the heart.* Berkeley: University of California Press.

Blackhall, L. J., Murphy, S. T., Frank, G., Michel, V., & Azen, S. (1995). Ethnicity and attitudes toward patient autonomy. *Journal of the American Medical Association, 274,* 820–825.

Blease, C. (2015). Authorized concealment and authorized deception: Well-intended secrets are likely to induce nocebo effects. *The American Journal of Bioethics, 15,* 23–25.

Boulware, L. E., Cooper, L. A., Ratner, L. E., LaVeist, T. A., & Powe, N. R. (2003). Race and trust in the health care system. *Public Health Reports, 118,* 358–365.

Boulware, L. E., Ratner, L. E., Cooper, L. A., Sosa, J. A., LaVeist, T. A., & Powe, N. B. (2002). Understanding disparities in donor behavior: Race and gender

differences in willingness to donate blood and cadaveric organs. *Medical Care, 40*(2), 85–95.

Bowen, O. R. (1954, February). Should patients be told when diagnosis is cancer? *Time, 15*, 71–72.

Buckman, R. (1984). Breaking bad news: Why is it still so difficult? *British Medical Journal, 288*, 1587–1589.

Callahan, D. (1992). When self-determination runs amok. *Hastings Center Report, 22*(2), 52–55.

Candib, L. (2002). Truth telling and advance planning at the end of life: Problems with autonomy in a multicultural world. *Families, Systems, and Health, 20*, 213–228.

Carrese, J. A., & Rhodes, L. A. (1995). Western bioethics on the Navajo reservation. *Journal of the American Medical Association, 274*, 826–829.

Chester v. Afshar. (2005). 1 AC 134, para 50, per Lord Hope.

Collins, J. (1927). Should doctors tell the truth. *Harper's, 155*, 320–326.

Edelstein, L. (1967). The Hippocratic Oath: Text, translation and interpretation. In O. Temkin & C. L. Temkin (Eds.), *Ancient medicine: Selected papers of Ludwig Edelstein* (pp. 3–64). Baltimore: Johns Hopkins University Press.

Eisenberg, D. J. (1986). *The change in physicians' attitudes towards truth-telling and to dying patients: 1920s–1980s.* D. J. Eisenberg Publisher.

Epstein, R. M., Korones, D., & Quill, T. E. (2010). Withholding information from patients—When less is more. *The New England Journal of Medicine, 362*, 380–381.

Epstein, R. M., & Peters, E. (2009). Beyond information: exploring patients' preferences. *Journal of the American Medical Association, 302*, 195–197.

Everett, J. P., Walters, C. A., Stottlemyer, D. L., Knight, C. A., Oppenberg, A. A., & Orr, R. D. (2011). To lie or not to lie: Resident physician attitudes about the use of deception in clinical practice. *Journal of Medical Ethics, 37*(6), 333–338.

Faden, R., Beauchamp, T., & King, N. M. P. (1986). *A theory and history of informed consent.* New York: Oxford University Press.

Feagin, J., & Bennefield, Z. (2014). Systemic racism and US health care. *Social Science and Medicine, 103*, 7–14.

Fitts, W. T., & Ravdin, I. S. (1953). What Philadelphia physicians tell patients with cancer. *Journal of the American Medical Association, 153*(10), 901–904.

Fitzgerald, C. J., & Danner, K. M. (2014). Malingering. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 641–642). Thousand Oaks: Sage.

Fox, R. C., & Swazey, J. P. (2008). *Observing bioethics.* New York: Oxford University Press.

Frank, G., Blackhall, L. J., Michel, V., Murphy, S. T., Azen, S. P., & Park, K. (1998). A discourse of relationships in bioethics: Patient autonomy and end-of-life decision making among elderly Korean Americans. *Medical Anthropology Quarterly, 12*(4), 403–423.

Garcia-Preto, N. (2016). Latino families: An overview. In M. McGoldrick, J. Giordano, & N. Garcia-Preto (Eds.), *Ethnicity and family therapy* (3rd ed., pp. 153–165). New York: The Guilford Press.

Gordon, E. J., & Daugherty, C. K. (2003). 'Hitting you over the head': Oncologists' disclosure of prognosis to advanced cancer patients. *Bioethics, 17*(2), 142–168.

Hartmann, N. (1932). *Ethics Vol-I.* London: George Allen and Unwin Ltd.

Henderson, L. J. (1935). Physician and patient as a social system. *New England Journal of Medicine, 212*, 819–823.

Hooker, W. (1849). *Physician and patient, a practical view of medical ethics*. New York: Arno Press.

Hubbell, A. P. (2014). Medical and pharmaceutical industry. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 651–653). Thousand Oaks: Sage.

Johnston, C., & Holt, G. (2006). The legal and ethical implications of therapeutic privilege-is it ever justified to withhold treatment information from a competent patient? *Clinical Ethics, 1*(3), 146–151.

Jones, J. H. (1992). The Tuskegee legacy AIDS and the Black community. *The Hastings Center Report, 22*(6), 38–40.

Jonsen, A. R. (2007). The god squad and the origins of transplantation ethics and policy. *The Journal of Law, Medicine, and Ethics, 35,* 238–240.

Kant, I. (1936). *Lectures on ethics* (L. Infeld, Trans.). New York: Harper & Row.

Katz, J. (2002). *The silent world of doctor and patient*. Baltimore: Johns Hopkins University Press.

Kaufman, W. (1953, January). Should doctors tell the truth? *Coronet, 33,* 22–26.

Krisman-Scott, M. A. (2000). An historical analysis of disclosure of terminal status. *Journal of Nursing Scholarship, 32*(1), 47–52.

Kübler-Ross, E. (1969). *On death and dying*. New York: Scribner's.

Kübler-Ross, E., Wessler, S., & Avioli, L. V. (1972). On death and dying. *Journal of the American Medical Association, 221*(2), 174–179.

Lamont, E. B., & Christakis, N. A. (1999). Some elements of prognosis in terminal cancer. *Oncology-Huntington, 13*(8), 1165–1170.

Larkin, C., & Searight, H. R. (2014). A systematic review of cultural preferences for receiving medical "bad news" in the United States. *Health, 6*(16), 2162–2173.

MacKenzie, C. R. (2009). What would a good doctor do? Reflections on the ethics of medicine. *HSS Journal: The Musculoskeletal Journal of Hospital for Special Surgery, 5*(2), 196–199.

Maloney, F. G. (1954). Should a patient be told he has cancer? *Wisconsin Medical Journal, 53,* 541–544.

Markel, H. (2004). "I swear by Apollo"—On taking the Hippocratic Oath. *New England Journal of Medicine, 350,* 2026–2029.

Martin, S. (2015). *A short history of disease: Plagues, poxes, and civilizations*. Harpenden Herts, UK: Oldcastle Books.

McKinney, Jr., R. E., Beskow, L. M., Ford, D. E., Lantos, J. D., McCall, J., Patrick-Lake, B., …, & Weinfurt, K. (2015). Use of altered informed consent in pragmatic clinical research. *Clinical Trials, 12,* 494–502.

Meisel, A. (1979). The exceptions to the informed consent doctrine: Striking a balance between competing values in medical decision-making. *Wisconsin Law Review, 2,* 413–488.

Meisel, A., & Kuczewski, M. (1996). Legal and ethical myths about informed consent. *Archives of Internal Medicine, 156,* 2521–2526.

Mullan, R. (2016). *Mad to be normal*. New York: Routledge.

Natanson v. Kline, 350 P 2d 1093, 1104 (Kansas, 1960).

Nie, J. B. (2013). *Medical ethics in China: A transcultural interpretation*. London: Routledge.

Novack, D. H., Plumer, R., Smith, R. L., Ochitill, H., Morrow, G. R., & Bennett, J. M. (1979). Changes in physicians' attitudes toward telling the cancer patient. *Journal of the American Medical Association, 241,* 897–900.

Oken, D. (1961). What to tell cancer patients: A study of medical attitudes. *Journal of the American Medical Association, 175,* 1120–1128.

Owens, D. C. (2018). *Medical bondage: Race, gender and the origins of American gynecology.* Athens, GA: University of Georgia Press.

Page, K. (2012). The four principles: Can they be measured and do they predict ethical decision-making? *BMC Medical Ethics, 13,* 1–8.

Pellegrino, E. D., & Thomasma, D. C. (1987). The conflict between autonomy and beneficence in medical ethics: Proposal for a resolution. *Journal of Contemporary Health Law and Policy, 3,* 23–46.

Pellegrino, E. D., & Thomasma, D. C. (1993). *The virtues in medical practice.* New York: Oxford University Press.

Pence, G. (2014). *Medical ethics: Accounts of groundbreaking cases* (7th ed.). New York: McGraw Hill.

Pence, G. (2017). *Medical ethics: Accounts of groundbreaking cases* (8th ed.). New York: McGraw Hill.

Percival, T. (2014). *Medical ethics.* London: Cambridge University Press.

Plante, M. L. (1967). An analysis of informed consent. *Fordham Law Review, 36,* 639–672.

Politi, M. C., Lewis, C. L., & Frosch, D. L. (2012). Supporting shared decisions when clinical evidence is low. *Medical Care Research and Review, 70,* 113S–128S.

Quill, T. E., & Townsend, P. (1991). Bad news: Delivery, dialogue, and dilemmas. *Archives of Internal Medicine, 151*(3), 463–468.

Reibel v. Hughes. (1980). 2 S.C.R. 880.

Reverby, S. M. (2009). *Examining Tuskegee: The infamous syphilis study and its legacy.* Chapel Hill: University of North Carolina Press.

Rising, M. L. (2017). Truth telling as an element of culturally competent care at end of life. *Journal of Transcultural Nursing, 28*(1), 48–55.

Rosenbaum, M. E., Ferguson, K. J., & Lobas, J. G. (2004). Teaching medical students and residents skills for delivering bad news: A review of strategies. *Academic Medicine, 79*(2), 107–117.

Rothman, K. (1991). *Strangers at the bedside.* New York: Basic Books.

Salgo v. Leland Stanford Jr. Board of Trustees, 317 P.2d 170 (Cal. Ct. App 1957).

Schloendorff v. Society of New York Hospital, 105 NE 92 (NY 1914).

Searight, H. R. (1992). Assessing patient competence for medical decision making. *American Family Physician, 45*(2), 751–759.

Searight, H. R. (2016). Guantanamo, anorexia nervosa, and death row: What are the limits of autonomy? *Psyccritiques*, *61*(21).

Searight, H. R., & Barbarash, R. A. (1994). Informed consent: Clinical and legal issues in family practice. *Family Medicine, 26*(4), 244–249.

Searight, H. R., & Gafford, J. (2005a). Cultural diversity at the end of life: Issues and guidelines for family physicians. *American Family Physician, 71*(3), 515–522.

Searight, H. R., & Gafford, J. (2005b). "It's like playing with your destiny": Bosnian immigrants' views of advance directives and end-of-life decision-making. *Journal of Immigrant Health, 7*(3), 195–203.

Searight, H. R., & Miller, C. K. (1996). Remembering and interpreting informed consent: A qualitative study of drug trial participants. *The Journal of the American Board of Family Practice, 9*(1), 14–22.

Sisk, B., Frankel, R., Kodish, E., & Isaacson, J. H. (2016). The truth about truth-telling in American medicine: A brief history. *The Permanente Journal, 20*(3), 74–77.

Smith, D. H. (1988). Communication as a reflection of and a source for values in health. *Journal of Applied Communication Research, 16*(1), 29–38.

Smith, R. C. (2002). *Patient-centered interviewing: An evidence-based method.* Philadelphia: Lippincott, Williams & Wilkins.

Smith, T. J., Dow, L. A., Khatcheressian, J., & Lyckholm, L. J. (2010). Giving honest information to patients with advanced cancer maintains hope. *Oncology, 24*(6), 521–528.

Stearns, S. A. (2014). Context. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 215–216). Thousand Oaks: Sage.

Stewart, M. A. (1995). Effective physician-patient communication and health outcomes: A review. *CMAJ: Canadian Medical Association Journal, 152*(9), 1423–1433.

Theobald, P., & Wood, K. L. (2009). Communitarianism and multiculturalism in the academy. *Journal of Thought, 44,* 9–23.

Tomes, N. (2016). *Remaking the American patient: How Madison Avenue turned patients into consumers.* Chapel Hill: The University of North Carolina Press.

Truog, R. D., Brown, S. D., Browning, D., Hundert, E. M., Rider, E. A., Bell, S. K., & Meyer, E. C. (2015). Microethics: The ethics of everyday clinical practice. *Hastings Center Report, 45,* 11–17.

Walter, P. (1997). The doctrine of informed consent: To inform or not to inform. *St. John's Law Review, 71,* 543–590.

Washington, H. A. (2006). *Medical apartheid: The dark history of medical experimentation on Black Americans from colonial times to the present.* New York: Doubleday Books.

Weir, R. (1980). Truthtelling in medicine. *Perspectives in Biology and Medicine, 24,* 95–112.

Yang, J., Scarfe, W. C., & Angelopoulos, C. (2018). Incidental findings on CBCT. In W. C. Scarfe & C. Angelopoulos (Eds.), *Maxillofacial cone beam computed tomography* (pp. 553–589). New York: Springer.

Yost, E. (1936, January). A patient wants to know. *Scribner's Magazine, 99,* 47–49.

Zhao, W. (2015). A Confucian worldview and family-based informed consent. A case of concealing illness from the patient in China. In F. Ruiping (Ed.), *Family-oriented informed consent: East Asian & American perspectives* (pp. 231–244). Dordrecht: Springer.

Deterring Deception: Approaches to Maximize Ethical Behavior in Social Interactions and Organizations

Lyn M. Van Swol, Evan Polman and Hangsan Paul Ahn

Through extensive research, researchers have found that both laypeople and professionals have a difficult time detecting deception. Rates of accurately detecting detection hover around chance levels (Aamodt & Custer, 2006; Bond & DePaulo, 2006). Further, nonverbal and demeanor cues that people assume to be useful for detecting deception often have little to no relationship to actual deceptive behavior (Levine et al., 2011; Van Swol & Braun, 2014; Vrij, 2008). Although third-party verification (verifying the consistency of a person's story) is more effective than relying on demeanor and nonverbal cues (Park, Levine, McCornack, Morrison, & Ferrerra, 2002), this method is time-consuming and requires access to extra information to the potentially deceptive message itself. What is more, interrogating questions that may be required to verify a person's story may be insulting and could damage the relationship between the sender and receiver of the message. These techniques may be appropriate in a law enforcement setting but not in a typical work setting that attempts to create positive organizational identity and foster citizenship among employees; or in a negotiation setting, in which parties want to establish a level of trust and cooperation. Given these problems, the best recourse in most situations may be to set

L. M. Van Swol (✉) · H. P. Ahn
Department of Communication Arts, University of Wisconsin-Madison, Madison, WI, USA
e-mail: vanswol@wisc.edu

E. Polman
Department of Marketing, University of Wisconsin-Madison, Madison, WI, USA

up situations to deter deception *outright*. Surveillance is one option that organizations can and do use, but surveillance is expensive and communicates a lack of trust (Cialdini, Petrova, & Goldstein, 2004; Van Swol, 2003) and can undermine intrinsic motivation to behave ethically (Tenbrunsel & Messick, 1999). Another option is to design situations to nudge people to behave more honestly.

A simple rational model of decision-making would predict that deception occurs when the benefits outweigh the risks (Becker, 1968; Grolleau, Kocher, & Sutan, 2014). Yet, even when risks of detection are minimal or even completely eliminated, people often do not cheat and lie as much as possible. People may assume that dishonesty is due to internal differences in morality and that some people do not cheat and deceive because they are just more honest or have a stronger moral compass than others (Moore & Gino, 2013). However, situational cues and social influences often have strong effects on a person's level of deception. Social norms, ease of rationalization, and salience of the unethicality of the behavior all play a role in increasing or decreasing deceptive behavior beyond what would be predicted by just weighing costs and benefits. These processes can work to make the ethical and moral implications of one's behavior more or less salient. When the moral implications of a deceptive action are out of awareness, then deceptive behavior is more likely (Butterfield, Treviño, & Weaver, 2000). In this chapter, we review ways that ethical behavior can be nudged into awareness in social interactions and organizations. We review five nudges that could affect deceptive behavior, including the framing of message or situation, social norms on what is deceptive or unethical, moral licensing, the salience of the deceptive action, and perceptions of fairness and justice.

Framing

One factor that can nudge honest or dishonest behavior is the framing of an option involving a resource. Research on prospect theory (Kahneman & Tversky, 1979) defined the concept of loss aversion in decision-making. Rather than evaluating options on the basis of total wealth possible from an outcome, options are evaluated on the basis of changes in wealth from a determined reference point. People are sensitive to gains and losses from a predetermined reference point and behave differently if they perceive the option as a possible gain or possible loss. Consistently, research has found that people are more risk-averse when they perceive an option as a gain, and more risk-seeking when they perceive the option as a loss (Kahneman & Tversky, 1984). Therefore, framing options as gains or losses from a reference point can affect risky behavior, such that people will prefer a greater risk to avoid a loss than pursue a gain. Identical situations can be approached quite differently then, whether they are framed as a gain or a loss. Further, losses often loom larger in people's decision-making. The prospect of a loss often motivates behavior more than the prospect of a gain. Because deception is often a risky behavior, it can occur more when situations are framed as losses than gains.

Accumulating research has found that frames can affect deceptive and selfish behavior. For example, Grolleau et al. (2014) found that participants, especially men, are more likely to cheat on a math-matrix solving task when they are in a loss frame than a gain frame. Rajgopal and White (2015) found that New York City cab drivers are more likely to overcharge and cheat customers when they perceive that they are going to make less money on a ride due to regulatory restrictions and will end up "in the hole" in comparison with a comparable ride to another location. Research on social dilemmas has found that participants are more selfish if they have to give up resources (called a "give some" game) than if they can take resources (called a "take some" game) (Brewer & Kramer, 1986; Neale & Bazerman, 1985). Consistent with this view, Cameron and Miller (2009) found that participants were more likely to cheat on an anagram solving task in a loss frame than a gain frame. And Kern and Churgh (2009) found that participants cheated more in scenarios when they were in a loss frame than a gain frame. Moreover, they found evidence that the effect of frames on ethical behavior is often a rapid and automatic process. Finally, Robben et al. (1990) found that people cheated more on their taxes when facing a balance due (loss) than a refund (gain). Thus, a gain frame can act to nudge more ethical and less deceptive behavior.

Another type of framing that can affect deceptive behavior involves regulatory focus. Regulatory focus theory differentiates between whether one is concerned with positive outcomes (called a promotion focus) or negative outcomes (called a prevention focus). A promotion focus emphasizes approaching goals, attainment, eagerness, and aspirations. A prevention focus emphasizes avoiding negative outcomes, vigilance, and upholding responsibilities (Higgins, 1997). Gino and Margolis (2011) hypothesized that a promotion focus leads to more risk-seeking behavior because a person is eager to take chances to attain a positive outcome and that a prevention focus leads to more caution and vigilance and thus less risk-taking. Because unethical behavior is inherently risky, they hypothesized and found more cheating among participants primed with a promotion focus. In a similar vein, other research has found that people who are strongly focused on winning, which is consistent with a promotion focus, are more likely to cheat in a prisoner's dilemma (Schweitzer, DeChurch, & Gibson, 2005). Solgos (2016) also found that a promotion focus increased cheating, possibly because participants with a promotion focus were eager to take chances. In addition, a prime reason that students cheat in school (Simkin & McLeod, 2010), and bankers engage in illegal transactions (Tenbrunsel & Thomas, 2015), is in their words "to get ahead" and reach their goal. Efforts to reach profit-oriented goals may be especially conducive to deceptive and unethical behavior in organizations as pursuit of goals may lead ethical concerns to be sidelined (Moore & Gino, 2013; Ordóñez, Schweitzer, Galinsky, & Bazerman, 2009; Wolfe, 1988). Therefore, loss or promotion frames which can encourage risk, especially when pursuing goals, may inadvertently also encourage deception.

SOCIAL NORMS

Another reason that people may deceive more in a loss frame is that there is a perception that deception and cheating are more socially acceptable if someone is trying to recover a loss (Kahneman, Knetsch, & Thaler, 1986). Social norms are important in terms of both nudging honest and encouraging deceptive behavior. What others do guides our own behavior, and generally speaking, people want to behave in ways that are socially approved (Cialdini, Kallgren, & Reno, 1991). For example, research on bystander intervention has found that people are less likely to intervene and help when others ignored a person seeming to suffer an emergency, thus creating a social norm toward non-intervention (Latané & Darley, 1970).

People have a conflicting desire between reaping the benefits of deception and cheating, and viewing themselves as good and moral people with high integrity (Ariely, 2012). Strong social norms and societal disapproval against certain types of deception may make the deception more salient and unpalatable due to its implications for one's integrity, especially as viewed by other people. For example, people are less likely to cheat someone who is visually handicapped than someone with normal range vision (Maharabani, 2007), even though chances of detection are lessened, possibly due to strong social norms against taking advantage of another person's handicap to increase one's gains. Other research has found that having participants read the Ten Commandments and swear on a Bible or reminding participants of an honor code essentially eliminated deceptive behavior, even among self-identified atheists, possibly because it highlighted social and religious norms against deception (Mazar, Amir, & Ariely, 2008). Social norms are more powerful when social information is coming from similar others or in-group members (Gino, Ayal, & Ariely, 2009; Moore & Gino, 2013). For example, individuals reported fewer tax deductions, which could be an indicator of less deception, when they received information that similar others complied with the tax code (Wenzel, 2005), but information about others' compliance did not affect behavior when individuals did not identify with those who were following the norm.

Social norms, however, are not always positive; cultures can develop norms that can encourage deception and cheating through the development of local norms, especially when cheating is highly profitable. MacLean (2001) documented how deceptive and profitable sales practices became normative at an insurance company, and the normalization of unethical and deceptive cultures in organizations is well documented (Ashforth & Anand, 2003; Toffler, 2003). Blatant cheaters can promote unethical behavior through a highly salient role model (Gino et al., 2009). If people feel psychologically close to someone engaging in unethical behavior, they often excuse the behavior and are more likely to engage in the behavior themselves (Gino & Galinsky, 2012). People also try to enlist others to be involved to some degree in their own deceptive activities as a way to justify the behavior and make it appear

more socially acceptable (Mazar, Shampanier, & Ariely, 2011). Therefore, creating social norms that discourage deception, especially when similar others follow the norms, and making the norms prominent in situations where the temptation to deceive is strong may help reduce deception. Societies and organizations should also avoid having blatant cheaters in highly public political or executive offices, as it sets a strong social norm for everyone else (Ariely, 2012), and groups and organizations should promote positive and visible in-group role models to foster honest behavior (Moore & Gino, 2013).

Moral Licensing

Just as controls of deception, like framing and social norms, can be used to decrease deception so too can such controls inadvertly foster deception. This type of backlash is called *moral licensing*, and it describes that people treat their moral and immoral behaviors like currency—as though good, moral behaviors (like donating to a colleague's charity run) are thought to accumulate and can thereupon be "exchanged" for the occasional misdeed (like canceling plans with a colleague at the last minute). Among the first papers to show evidence of moral licensing, Monin and Miller (2001) found in a set of studies that providing people with an opportunity to express non-sexist attitudes subsequently led them to behave in a relatively prejudiced manner—as though their initial non-sexist behavior inoculated them from appearing prejudiced. Similarly, Effron, Cameron, and Monin (2009) found that expressing support for an African-American candidate led participants to demonstrate racist behaviors by favoring Whites at the expense of Blacks. Importantly, people don't need to *do* anything before deciding that they have earned the right to relax their behavior. Rather, a pattern of moral licensing can ensue even after a decision to forgo a questionably dubious or shady choice (Effron, Miller, & Monin, 2012). For example, people who pass up on ordering the greasy cheeseburger in favor of ordering something healthier may feel relatively proud of their choice—e.g., they may feel emboldened because they bested temptation and made the "right" choice. However, this particular feeling and behavior has been found to subsequently lead people to feel like they have earned a guilty pleasure; that is, because of their good behavior, people feel like they have earned the right to indulge (Kotabe, Righetti, & Hofmann, 2014). As these examples point out, individuals' perceived moral standing is flexible and can change as moral (and immoral) behaviors accumulate (Miller & Effron, 2010). Accordingly, after engaging in a commendable, generous behavior, people are sometimes more willing to let their good behavior slip, as their moral balance has tipped in favor of a positive (in contrast to negative) self-image.

As evidenced, when people acquire moral capital, they establish themselves (and to others) as kind, generous, or courteous human beings—like model citizens—which is subsequently viewed as justification to relax their behavior, by taking "more liberties" and letting their behavior fall a little south

of appropriate (Khan & Dhar, 2006). Relatedly, recent evidence has found that people feel like they have the right to behave however they please on their birthdays, as if this entitles one to a free pass for moral transgressions on that day (Moore & Pierce, 2016). This shows that people "invent" reasons to commit an occasional misdeed. For example, people are more likely to cheat or steal if their misbehavior benefits other people (Gino, Ayal, & Ariely, 2014). Viewed this way, licensing can be seen as a defense mechanism. For example, research has documented that sometimes the people who are the harshest judges of others' misbehavior are people who themselves have likewise misbehaved (Barkan, Ayal, Gino, & Ariely, 2012). Not unlike "the pot calling the kettle black," this behavior was evidenced in a famous case when a former Massachusetts Institute of Technology (MIT) dean of admissions was known for her especially severe attitude toward students who lied on their application—an attitude she might have held because, as was discovered later, she had lied on her own resume, claiming bachelor's and master's degrees she did not actually have (Lewin, 2007). Quite possibly, the dean's harsh attitude in this example is a defensive behavior, in that appearing strongly against lying inoculated her from suspicions others might potentially raise about her own lying. In a similar manner, people sometimes show extensive outrage at transgressions leveled against others, because appearing infuriated in this way can be seen as a strategy of lessening guilt over one's own moral failings (Rothschild & Keefer, 2017).

In most cases of moral licensing, people tend to behave in dishonest, maligned ways yet still perceive themselves as honest, benevolent people (Jordan, Mullen, & Murnighan, 2011; Mazar et al., 2008). In part, this is because people tend to forget their misdeeds (Sezer, Gino, & Bazerman, 2015) and exaggerate their good ones (Effron, 2014). This behavior is problematic from the perspective of deterrence, because a practical strategy for deterring deception is reminding people of their deception (Ayal, Gino, Barkan, & Ariely, 2015), yet as evidenced, people may have difficulty with generating examples, believing they are more like saints than sinners (Epley & Dunning, 2000). This self-aggrandizing belief has been shown to be particularly poignant. As an example, Elliott Spitzer, the former governor of New York, was eventually punished for privately patronizing prostitutes despite publicly working to punish people who did the same thing (Hakim & Rashbaum, 2008). How can someone sleep with prostitutes and persecute others who do the same? Put differently, how can someone clench to two opposing beliefs or behaviors? The answer lies in ethical dissonance, and specifically in an area of moral hypocrisy (Barkan, Ayal, & Ariely, 2015; Monin & Merrit, 2011), which describes that people hold different standards for themselves, believing that transgressions committed by the self are less bad than the same transgressions committed by others (Docan-Morgan & Docan, 2007). More the rule than the exception, people hold opposing beliefs of their deceptive behavior, but on the bright side, research has found some ways that reduce this type of

hypocrisy. For example, moral hypocrisy is decreased by feelings of guilt (Polman & Ruttan, 2012), feelings of gratitude (Tong & Yang, 2011), and cognitive constraint (Valdesolo & DeSteno, 2008).

Paradoxically, the problem with moral licensing is that deterrences to deception may inadvertently lead to other kinds of deception. Which is to say that when people feel like they have successfully evaded a potentially selfish or deceptive choice, they may "treat" themselves with some leniency on a future transgression. This form of self-regulation is somewhat ambiguous—it involves people distorting assessments and the severity of their misdoings. For this reason, people who are especially creative (i.e., well-equipped to reinterpret behavior in new ways) are also more likely to engage in unethical behaviors (Gino & Ariely, 2012; Gino & Wiltermuth, 2014), to the extent that creative types may not feel like their transgressions are transgressions at all, but in fact, prosocial behaviors (Vincent & Polman, 2016). Suffice to say, people are scarily adept at reinterpreting immoral behavior. In another example, people ordinarily disparage sweatshop labor, but upon discovering that their favorite jeans were made under poor working conditions, people warmed up to the idea of sweatshop labor, even defending it (Paharia, Vohs, & Deshpandé, 2013). In a similar sort of vein, people are also defensive of others' misbehaviors, for example, by giving close others a pass for their transgressions (Valdesolo & DeSteno, 2007) and by downplaying the deception of respected others (Polman, Pettit, & Wiesenfeld, 2013).

Needless to say, solutions to deterring people's deception is particularly challenging when considering the effects of moral licensing. As research has found, swaying people away from making a deceptive choice could backfire, insofar as people could consider this non-choice as fodder (as justification) for making a deceptive or immoral choice in the near future. At least in some cases, deterrence is a double-edged sword—it cuts deception in two ways, by both preventing deception and fostering it. Viewed this way, deterring deception is like taking care of a plant: pruning its leaves will make it smaller, but in the long term, pruning will make it even larger. Analogous to taking care of a plant, curbing deception will lessen it in the short term, but not without potentially seeding long-term effects that encourage deception.

That said, some solutions to licensing have been documented. One particular strategy based on misattribution is especially encouraging (Gallier, Reif, & Römer, 2017). As an example, people tend not to license and thus commit a misdeed if they were paid for their initial, positive behavior (Clot, Grolleau, & Ibanez, 2013). This strategy is consistent with previous work on attribution theory that finds that external incentives for performing prosocial behavior crowd out any other internal incentives (Kelley, 1973), as though prosocial behavior that is not intrinsically motivated loses its currency in moral licensing. In further support, Khan and Dhar (2006) compared a voluntary prosocial behavior with an involuntary one (e.g., performing community service as a result of a driving violation). Their results show that people license in

the former case, when the prosocial behavior was voluntary, but significantly less so in the latter case. Evidently, in order for licensing effects to emerge, the initial behavior has to be intrinsically motivated, such that if people have an alternative, external reason for their positive behavior then licensing disappears. Entirely consistent with this view, Gneezy, Imas, Nelson, Brown, and Norton (2012) found that only moral actions that contain real costs have signaling power, precisely the type of power that can foster licensing.

Another related strategy is, essentially, making people feel less special. As noted, creative people tend to be more dishonest compared to people who are less creative (Vincent & Polman, 2016). However, the relation between creativity and dishonesty disappears when creative people no longer think that their creativity is rare or special (Vincent & Kouchaki, 2016). For example, when participants were led to believe that they "performed very well" on a task in which "creativity is common" they lied less on a subsequent task compared to participants who "performed very well" on a task in which "creativity is rare" (Vincent & Kouchaki, 2016). In this vein, people have a tendency to feel morally superior to others (Tappin & McKay, 2016); however, if an intervention were to humble people and provide them with information that suggests they are more similar than superior to others, then they may license less, feeling like they have not acquired sufficient moral capital to spend on misdeeds.

Finally, another strategy for deterring licensing is, counter-intuitively, to allow people to deceive, at least a little bit. The self-regulation model of exchanging good behaviors for bad ones works both ways, such that people also exchange bad behaviors for good behaviors (Zhong, Ku, Lount, & Murnighan, 2010). In complement to moral licensing, this is called *moral cleansing*, which describes that behaving immorally has a negative influence on perceptions of self-worth, which leads people to engage in moral behavior in order to regain some of that lost worth (Sachdeva, Iliev, & Medin, 2009). As an example in practice, it might be effective to indulge in a small unhealthy treat, even if one has a dieting goal and is trying to avoid such treats, because the small treat provides the fodder and motivation to go to the gym and eat extra healthy for the rest of the day.

SALIENCE AND UNAMBIGUITY OF DECEPTIVE ACTION

Despite people's penchant for resolving ethical dissonance, and for moral hypocrisy, people find it challenging to *blatantly* engage in deception or cheat a lot and still feel good about themselves due to the need to preserve a sense of integrity (Mazar & Ariely, 2006). While factors that reduce the salience of the deceptive action can inadvertently help individuals rationalize their deceptive behavior and reduce the negative implications for their integrity, it follows that increasing the salience of the ethicality of one's actions can reduce cheating by making it harder to rationalize deception. For example, participants who signed their name to verify the truthfulness of their answers before filling out a form (i.e., reporting odometer or travel reimbursements) were

less deceptive than participants who signed after filling out the form (Shu, Mazar, Gino, Ariely, & Bazerman, 2012). This is because an initial signature makes the ethicality of one's answers more salient. Verbal or situational cues that unambiguously label a behavior as deceptive generally reduce deception and cheating, as it makes it harder to rationalize a deceptive behavior that is clearly labeled as such and still preserve one's sense of integrity (Gino & Ariely, 2012; Gino et al., 2009). Which is to say that people will usually behave ethically if the choice is very clear (Butterfield et al., 2000).

Actions that can decrease the psychological distance between the self and deceptive act can also increase the salience of deception and make disambiguation harder. For example, in one study in which participants were paid in tokens for correct answers on a test, participants were more likely to cheat and steal tokens (that could be exchanged for money) compared to participants who were paid in actual money. This is because stealing cash increases the salience that one is cheating; however, stealing tokens (that can be readily exchanged for cash) feels less like stealing (Mazar et al., 2008). In a similar manner, golfers are more likely to cheat by moving the ball with their club than their hand, presumably because the club provides distance between oneself and the deceptive behavior (McKenzie, 2009). When there is a larger temporal lag between a deceptive act and its outcomes, deception is likely to be less salient and more likely to occur. In more support, McKenzie (2009) also found that golfers are more likely to allow themselves a penalty-free extra shot on the first hole of nine holes, when the consequences of the extra shot on the final score are less known than on the ninth hole. Likewise, use of sanitizing language and euphemisms also help create psychological distance from the consequences and implications of a deceptive act (Moore & Gino, 2013). For example, Enron traders relabeled fraud and deception as "arbitrage opportunities" (McLean & Elkind, 2004).

Understanding the ethical implications of dishonest behavior may take more conscious processing and time. Research has found that under the pressure of time, people are more likely to lie for money on a dice-rolling task, presumably because self-interest may be more of an automatic, non-volitional process, such that refraining from cheating for personal gain requires self-control and more conscious processing (Shalvi, Eldar, & Bereby-Meyer, 2012). In line with a relation between cheating and automatic processes, other research has found that cheating and deception are more likely to occur later in the day when people have fewer cognitive resources to resist or understand the salience of deceptive behavior (Kouchaki & Smith, 2014). Indeed, people are more likely to engage in deception when they are cognitively depleted (Mead, Baumeister, Gino, Schweitzer, & Ariely, 2009). In all, people may need time to understand the ethical implications of their selfish behavior. In addition, people may need to build up their level of self-control (Muraven, 2010) so that they are less likely to be depleted and tempted with deception.

When deceptive behavior can benefit others as well as oneself, it reduces the salience of deception and is easier to justify the deception and preserve

one's integrity. Thus, rather than focusing on the unethical implications of the deceptive action, one can focus on helping others (albeit through deception and cheating) and mitigate perceptions of immorality or greed (Gino, Ayal, & Ariely, 2014; Wiltermuth, 2011). This is problematic because deceptive behavior in organizations can often be easily justified as helping the organization's bottom line (Hildreth, Gino, & Bazerman, 2016; Umphress & Bingham, 2011).

In conclusion, in order to reduce deceptive and unethical behavior, it is necessary to take steps to highlight the moral and ethical implications of decisions. Individuals can be very creative at ethical fading, especially when deception is beneficial and profitable, but if the choice is made clear and unambiguous and the moral implications highlighted, most individuals will strive to be honest (Schweitzer & Hsee, 2002; Tenbrunsel & Messick, 2004). Thus, factors that increase self-awareness and mindfulness can help individuals connect the dots between their behavior and actions (Moore & Gino, 2013). For example, research has found that having individuals view themselves in a mirror reduced unethical behavior and self-preferential treatment (Batson, Thompson, Seuferling, Whitney, & Strongman, 1999; Diener & Wallbom, 1976). Highlighting codes of conduct may also help highlight that a behavior is deceptive (Treviño, Butterfield, & McCabe, 1998).

FAIRNESS AND RESTORING JUSTICE

When people perceive that they are entitled to more either through their own labor or because of characteristics of others, people will engage in more deception. People who worked to earn money, rather than receive the money through a random process, are more likely to steal or lie to keep more of the money (Gravert, 2013; Van Swol & Braun, 2014). When another person has been rude, people are more likely to cheat and not report a monetary error that benefits them and hurts the other person (Gneezy & Ariely, 2010). The rudeness of the other person may increase the sense of entitlement to cheat them out of money. Categorization of another person as an out-group member can increase the sense of entitlement to cheat and deceive the out-group member because people may dehumanize and ignore the distress of out-group members (Moore & Gino, 2013; Opotow, 1990).

People can also feel entitled to resources gained through deception and cheating when they perceive they are receiving inequitable outcomes or are mistreated (Gino & Pierce, 2009; Greenberg, 1990). Workers in the fast-food industry justify theft and deviancy through mistreatment by employers (Hollinger, Slora, & Terris, 1992). Deception offers a way to claim resources to which employees feel entitled (Greenberg, 1990). In order to minimize deception due to entitlement, organizations must ensure that employees perceive outcomes as equitable and clearly communicate with employees to ensure procedural and distributive justice (Moore & Gino, 2013).

Applications and Further Research

Increases in ethical behavior benefit both the actors and organizations in which they work by creating work environments in which ethical behavior can be achieved with lower costs and less surveillance. Understanding that the environment and norms affect the willingness of people to cheat and deceive can help organizations, governments, and people design situations that maximize honest behavior. Given the above review, here we suggest five practical applications that could help deter deception in social interaction and organizational contexts. First, consider message framing and using a gain frame in situations involving resources where there may be a temptation to cheat. For example, emphasize what people can gain from good performance, rather than what they would lose for bad performance. Another consideration is whether to emphasize the benefits and positive outcomes (promotion focus) of an action or the negative outcomes (prevention focus) to be avoided. A promotion focus may lead to more risky behaviors, which could include cheating and deception. Thus, a more cautious, prevention focus could foster more honesty. Second, creating strong norms toward honesty can deter deception, and positive norms and role models should be salient. For example, the leader of a country or organization should provide positive examples of following the law, paying taxes, and avoiding the taint of corruption. Third, people need to be mindful that measures taken to reduce deception could inadvertently increase later deception through moral licensing. Thus, having people attribute honesty to external sources, rather than their own good nature, or letting people deceive a little may be measures to prevent people from justifying deception because of their past good deeds. For example, rewarding honest behavior may reduce later deception by providing an external attribution to the honest behavior that could reduce moral licensing. Fourth, when possible, deceptive actions should be unambiguously labeled as such to avoid easy rationalization. For example, contracts could specify at the outset what is considered unethical, or require a signature from a person that they are answering honestly, or the targets of the deceptive behavior could be made salient. Finally, organizations and people should work to ensure that people feel treated equitable, as deception and dishonesty may be used in organizations as a way to restore equity.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15,* 6–11.

Ariely, D. (2012). *The (honest) truth about dishonesty: How we lie to everyone—Especially ourselves.* New York: Harper.

Ashforth, B. E., & Anand, V. (2003). The normalization of corruption in organizations. *Research in Organizational Behavior, 25,* 1–52.

Ayal, S., Gino, F., Barkan, R., & Ariely, D. (2015). Three principles to REVISE people's unethical behavior. *Perspectives on Psychological Science, 10*(6), 738–741.

Barkan, R., Ayal, S., & Ariely, D. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology, 6*, 157–161. https://doi.org/10.1016/j.copsyc.2015.08.001.

Barkan, R., Ayal, S., Gino, F., & Ariely, D. (2012). The pot calling the kettle black: Distancing response to ethical dissonance. *Journal of Experimental Psychology: General, 141*(4), 757–773.

Batson, C. D., Thompson, E. R., Seuferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology, 77,* 525–537.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The Economic Dimensions of Crime* (pp. 13–68). London, UK: Palgrave Macmillan.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234.

Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size and decision framing. *Journal of Personality and Social Psychology, 50,* 543–549.

Butterfield, K. D., Treviño, L. K., & Weaver, G. R. (2000). Moral awareness in business organizations: Influence of issue-related and social context factors. *Human Relations, 53*, 981–1018.

Cameron, J. S., & Miller, D. T. (2009). Different ethical standards in gain versus loss frames. In D. de Cremer (Ed.), *Psychological perspectives on ethical behavior and decision making* (pp. 91–106). Charlotte, NC: Information Age Publishing.

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology, 24,* 201–234.

Cialdini, R. B., Petrova, P. K., & Goldstein, N. J. (2004). The hidden costs of organizational dishonesty. *MIT Sloan Management Review, 45,* 67–73.

Clot, S., Grolleau, G., & Ibanez, L. (2013). Self-licensing and financial rewards: Is morality for sale? *Economics Bulletin, 33,* 2298–2306.

Diener, E., & Wallbom, M. (1976). Effects of self-awareness on antinormative behavior. *Journal of Research in Personality, 10,* 107–111.

Docan-Morgan, T., & Docan, C. A. (2007). Infidelity on the Internet: Double standards and the differing views of women and men. *Communication Quarterly, 55,* 317–342.

Effron, D. A. (2014). Making mountains of morality from molehills of virtue threat causes people to overestimate their moral credentials. *Personality and Social Psychology Bulletin, 40*(8), 972–985.

Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring whites. *Journal of Experimental Social Psychology, 45*(3), 590–593.

Effron, D. A., Miller, D. T., & Monin, B. (2012). Inventing racist roads not taken: The licensing effect of immoral counterfactual behaviors. *Journal of Personality and Social Psychology, 103*(6), 916–932.

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self-or social prediction? *Journal of Personality and Social Psychology, 79*(6), 861–875.

Gallier, C., Reif, C., & Römer, D. (2017). Repeated pro-social behavior in the presence of economic interventions. *Journal of Behavioral and Experimental Economics, 69,* 18–28.

Gino, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology, 102,* 445–459.

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science, 20,* 393–398.

Gino, F., Ayal, S., & Ariely, D. (2014). Self-serving altruism: The lure of unethical actions that benefit others. *Journal of Economic Behavior & Organization, 93,* 285–292.

Gino, F., & Galinsky, A. D. (2012). Vicarious dishonesty: When psychological closeness creates distance from one's own moral compass. *Organizational Behavior and Human Decision Processes, 119,* 15–26.

Gino, F., & Margolis, J. D. (2011). Bringing ethics into focus: How regulatory focus and risk preferences influence (un)ethical behavior. *Organizational Behavior and Human Decision Processes, 115,* 145–156.

Gino, F., & Pierce, L. (2009). Dishonesty in the name of equity. *Psychological Science, 20,* 1153–1160.

Gino, F., & Wiltermuth, S. S. (2014). Evil genius? How dishonesty can lead to greater creativity. *Psychological Science, 25*(4), 973–981.

Gneezy, A., & Ariely, D. (2010). *Don't get mad, get even: On consumers' revenge* (Working Paper). Duke University.

Gneezy, A., Imas, A., Nelson, L. D., Brown, A., & Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behavior. *Management Science, 58*(1), 179–187.

Gravert, C. (2013). How luck and performance affect stealing. *Journal of Economic Behavior & Organization, 93,* 301–304.

Greenberg, J. (1990). Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts. *Journal of Applied Psychology, 75,* 561–568.

Grolleau, G., Kocher, M. G., & Sutan, A. (2014). *Cheating and loss aversion: Do people lie more to avoid a loss?* (CESifo Working Paper No. 4965).

Hakim, D., & Rashbaum, W. K. (2008, March 10). Spitzer is linked to prostitution ring. *New York Times.* Retrieved online on February 16, 2018 at http://www.nytimes.com/2008/03/10/nyregion/10cnd-spitzer.html.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 55,* 1280–1300.

Hildreth, J. A. D., Gino, F., & Bazerman, M. (2016). Blind loyalty? When group loyalty makes us see evil or engage in it. *Organizational Behavior and Human Decision Processes, 132,* 16–36.

Hollinger, R. C., Slora, K. B., & Terris, W. (1992). Deviance in the fast-food restaurant: Correlates of employee theft, altruism, and counterproductivity. *Deviant Behavior, 13,* 155–184.

Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin, 37*(5), 701–713.

Kahneman, D., Knetsch, I., & Thaler, R. (1986). Fairness as a constraint on profit seeking. *American Economic Review, 76,* 728–741.

Kahneman, E., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291.

Kahneman, E., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39,* 341–350.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist, 28*(2), 107–128.

Kern, M. C., & Chugh, D. (2009). Bounded ethicality. *Psychological Science, 20,* 378–383.

Khan, U., & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research, 43*(2), 259–266.

Kouchaki, M., & Smith, I. H. (2014). The morning morality effect: The influence of time of day on ethical behavior. *Psychological Science, 25,* 95–102.

Kotabe, H. P., Righetti, F., & Hofmann, W. (2014). Affective forecasting in self-control. *Unpublished manuscript.*

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., et al. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research, 37,* 377–403.

Lewin, T. (2007, April 27). Dean at M.I.T. resigns, ending a 28-year lie. *New York Times.* Retrieved online on February 16, 2018 at http://www.nytimes.com/2007/04/27/us/27mit.html.

MacLean, T. (2001). Thick as thieves: A socially embedded model of rule breaking in organizations. *Business and Society, 40,* 167–196.

Maharabani, E. (2007). *Honesty and helping behavior: Testing situations involving temptation to cheat a blind person* (Master's dissertation). Ben-gurion University of the Negav.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633–644.

Mazar, N., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Journal of Public Policy & Marketing, 25,* 117–126.

Mazar, N., Shampanier, K., & Ariely, D. (2011). Probabilistic price promotions—When retailing and Las Vegas meet. *Management Science, 63*(1), 250–266. https://doi.org/10.1287/mnsc.2015.2328.

McKenzie, S. (2009). *Driven to cheat: A study on the drivers of dishonesty—Through the game of golf* (Undergraduate Honors Theses and Student paper, Duke University). Available from Duke space database.

McLean, B., & Elkind, P. (2004). *The smartest guys in the room: The amazing rise and scandalous fall of Enron.* New York, NY: Penguin.

Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E., & Ariely, D. (2009). Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology, 45*(3), 594–597.

Miller, D. T., & Effron, D. A. (2010). Psychological license: When it is needed and how it functions. In M. Zanna & J. Olson (Eds.), *Advances in experimental social psychology* (pp. 117–157). Stanford, CA: Elsevier.

Monin, B., & Merritt, A. C. (2011). Moral hypocrisy, moral inconsistency, and the struggle for moral integrity. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil. Herzliya series on personality and social psychology* (pp. 167–184). Washington, DC: American Psychological Association.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81*(1), 33–43.

Moore, C., & Gino, F. (2013). Ethically Adrift: How others pull our moral compass from true north, and how we can fix it. *Research in Organizational Behavior, 33,* 53–77.

Moore, C., & Pierce, L. (2016). Reactance to transgressors: Why authorities deliver harsher penalties when the social context elicits expectations of leniency. *Frontiers in Psychology, 7,* 550.

Muraven, M. (2010). Building self-control strength: Practicing self-control leads to improved self-control performance. *Journal of Experimental Social Psychology, 46,* 465–468.

Neale, M. A., & Bazerman, M. H. (1985). The effects of framing and negotiator overconfidence on bargaining behaviors and outcomes. *The Academy of Management Journal, 28,* 34–49.

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues, 46,* 1–20.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives, 23,* 6–16.

Paharia, N., Vohs, K. D., & Deshpandé, R. (2013). Sweatshop labor is wrong unless the shoes are cute: Cognition can both help and hurt moral motivated reasoning. *Organizational Behavior and Human Decision Processes, 121*(1), 81–88.

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrera, M. (2002). How people really detect lies. *Communication Monographs, 69,* 144–157.

Polman, E., Pettit, N. C., & Wiesenfeld, B. M. (2013). Effects of wrongdoer status on moral licensing. *Journal of Experimental Social Psychology, 49*(4), 614–623.

Polman, E., & Ruttan, R. L. (2012). Effects of anger, guilt, and envy on moral hypocrisy. *Personality and Social Psychology Bulletin, 38*(1), 129–139.

Rajgopal, S., & White, R. (2015). *Cheating when in the hole: The case of New York city taxis* (Working Paper). Emory University.

Robben, H. S. J., Webley, P., Weigel, R. H., Warneryd, K., Kinsey, K. A., Hessing, D. J., et al. (1990). Decision frame and opportunity as determinants of tax cheating: An international experiment. *Journal of Economic Psychology, 11,* 341–364.

Rothschild, Z. K., & Keefer, L. A. (2017). A cleansing fire: Moral outrage alleviates guilt and buffers threats to one's moral identity. *Motivation and Emotion, 41*(2), 209–229.

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners the paradox of moral self-regulation. *Psychological Science, 20*(4), 523–528.

Schweitzer, M. E., DeChurch, L. A., & Gibson, D. E. (2005). Conflict frames and the use of deception: Are competitive negotiators less ethical? *Journal of Applied Social Psychology, 35,* 2123–2149.

Schweitzer, M. E., & Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty, 25,* 185–201.

Sezer, O., Gino, F., & Bazerman, M. H. (2015). Ethical blind spots: Explaining unintentional unethical behavior. *Current Opinion in Psychology, 6,* 77–81.

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science, 23,* 1264–1270.

Shu, L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences of the United States of America, 109*(38), 15197–15200.

Simkin, M. G., & McLeod, A. (2010). Why do college students cheat? *Journal of Business Ethics, 94*(3), 441–453.

Solgos, J. T. (2016). *The effect of regulatory focus on ethical decision-making* (Doctoral dissertation). Ohio University.

Tappin, B. M., & McKay, R. T. (2016). The illusion of moral superiority. *Social Psychological and Personality Science, 8*, 1–9.

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly, 44*, 684–707.

Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self deception in unethical behaviour. *Social Justice Research, 17*, 223–236.

Tenbrunsel, A. E., & Thomas, J. (2015). *The street, the bull and the crisis: A survey of the US & UK financial services industry*. New York: Labaton Sucharow.

Toffler, B. L. (2003). *Final accounting: Ambition, greed, and the fall of Arthur Andersen*. New York: Random House.

Tong, E. M., & Yang, Z. (2011). Moral hypocrisy: Of proud and grateful people. *Social Psychological and Personality Science, 2*(2), 159–165.

Treviño, L. K., Butterfield, K. D., & McCabe, D. L. (1998). The ethical context in organizations: Influences on employee attitudes and behaviors. *Business Ethics Quarterly, 8*, 447–477.

Umphress, E. E., & Bingham, J. B. (2011). When employees do bad things for good reasons: Examining unethical pro-organizational behaviors. *Organization Science, 22*, 621–640.

Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy social groups and the flexibility of virtue. *Psychological Science, 18*(8), 689–690.

Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology, 44*(5), 1334–1338.

Van Swol, L. M. (2003). The effects of regulation on trust. *Basic and Applied Social Psychology, 25*, 221–233.

Van Swol, L. M., & Braun, M. T. (2014). Channel choice, justification of deception, and detection. *Journal of Communication, 64*, 1139–1159.

Vincent, L. C., & Kouchaki, M. (2016). Creative, rare, entitled, and dishonest: How commonality of creativity in one's group decreases an individual's entitlement and dishonesty. *Academy of Management Journal, 59*(4), 1451–1473.

Vincent, L. C., & Polman, E. (2016). When being creative frees us to be bad: Linking creativity with moral licensing. In J. W. van Prooijen & P. A. M. van Lange (Eds.), *Cheating, corruption, and concealment: The roots of dishonesty* (pp. 166–184). Cambridge, UK: Cambridge University Press.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, UK: Wiley.

Wenzel, M. (2005). Misperceptions of social norms about tax compliance: From theory to intervention. *Journal of Economic Psychology, 26*, 862–883.

Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes, 115*, 157–168.

Wolfe, D. (1988). Is there integrity in the bottomline? Managing obstacles to executive integrity. In S. Srivastva (Ed.), *Executive integrity: The search for high human values in organization life* (pp. 140–171). San Francisco: Jossey-Bass.

Zhong, C. B., Ku, G., Lount, R. B., & Murnighan, J. K. (2010). Compensatory ethics. *Journal of Business Ethics, 92*(3), 323–339.

Contexts of Deceptive Communication: Law Enforcement Interrogations

CHAPTER 36

Verbal Cues Fostering Perceptions of Credibility and Truth/Lie Detection

Ray Bull, Maureen van der Burgh and Coral Dando

Cues to Credibility

Around 30 years ago, Bell and Loftus (1988) innovatively found that the amount of detail provided by witnesses can have a substantial effect on (mock) jurors' decisions. Since then a considerable number of research studies have investigated which factors regarding people's accounts of events influence decisions about the likely truthfulness/falsehood of such accounts made by others (such as investigators/police, lawyers, and judges).

We might wonder if it really matters which aspects of accounts relate to the believed credibility of what people say. Well, in one crucial arena it seems to. All of us are aware that in several countries around the world people are horrendously persecuted and due to this they flee to seek asylum in another country (i.e., becoming a refugee). In the UK in 2015 the relevant government ministry updated its guidance on "Assessing credibility and refugees status" in which it was stated that the credibility of refugees'/asylum seekers' claims can, in part, be established if their accounts/statements are (among other things)—"...of sufficient detail and specificity," "internally consistent and coherent," and "plausible" (Home Office, 2015).

R. Bull (✉)
University of Derby, Derby, UK
e-mail: rhb10@leicester.ac.uk

M. van der Burgh
Diagnostic Expertise Centre, Amsterdam, The Netherlands

C. Dando
University of Westminster, London, UK

The guidance provided by the United Nations High Commissioner for Refugees also mentions consistency, coherence, and plausibility. In 2003, Kagan stated that credibility assessment is often the single most important step in determining whether people seeking protection as refugees can be granted asylum. He also pointed out that a large proportion of rejected/unsuccessful applications are due to their low credibility, partly because of insufficient detail being provided by the applicants. However, when this chapter's first author, Ray Bull, was asked (within the UK) to examine how well the interviews with asylum seeking applicants were conducted he found that the interviewers were reasonably skilled but seemed to "back off" from asking detail seeking questions/hearing about "shocking" detail of maltreatment. This small-scale study has not been published.

Beliefs About Useful Cues

Many dozens of research studies have focused on the cues that people believe to be useful to decipher credibility, truth, and lies. Many of these (believed) cues relate to "nonverbal behaviour" (e.g., eye gaze) but fewer relate to the actual content of what people say or to how they say it (e.g., speech pauses). This chapter will focus on content. One likely reason why research has repeatedly found the behavioral cues people believe in to be of very limited use in detecting truth/deception is that liars use our beliefs to trick us.

Content Cues to Credibility

Over 30 years ago, Riggio, Tucker, and Widaman (1987) found that judgments of communicators' believability were significantly influenced by their verbal fluency. Canter, Ioannou, Youngs, and Chungh (2016) showed their research participants real-life video recordings of people making appeals for information about a missing relative that were shown on TV; half of these appeals were indeed genuine but the other half were false, in that the person making the appeal had actually already murdered their family member. The research participants were asked to assess each appeal/appealer and it was found that the plausibility of what the appealers said was significantly associated with assessments of credibility. In a study of sexual assault investigators, Campbell, Menaker, and King (2015) examined what these professionals said they used to assess "victim" credibility. Consistency of statements, contradictions, and amount of detail were reported to be widely used cues. Similarly, in a study of police officers' and lay people's beliefs about cues to deception, Bogaard, Meijer, Vrij, and Merckelbach (2016) found that quantity of details, contradictions, coherence, and plausibility were deemed important. These types of verbal/speech cues to lies/truths were also believed to be important by people taking part in other studies (e.g., Masip, Bethencourt, Lucas, Sanchez-San Segundo, & Herrero, 2012).

But Which (Believed in) Cues Are Actually Valid?

Sweeney (2009) made the important point that although an account may be thought/assessed/believed to be credible, this does not necessarily mean that the account is true. To this, we could add that some skilled liars may know what makes an account appear credible. In terms of the actual content of what people say, Porter and Yuille (1996) found that only three of the possible (believed in) cues they examined actually discriminated between when students were lying or telling the truth, these cues being (i) amount of detail reported, (ii) coherence, and (iii) admissions of lack of memory. Klaver, Lee, and Hart (2007) noted a higher rate of speech disturbances in prisoners' lying accounts about a crime video recording shown to them. Sooniste, Granhag, Stromwall, and Vrij (2015) also found the amount of detail provided by students and lay people to discriminate between truth and lies in response to unanticipated questions about the planning of an event. Masip, Blandon-Gitlin, Martinez, Herrero, and Ibabe (2016) found that when students lied to questions (on two occasions) about committing a ("mock") crime their answers were less consistent than when truth-telling. Brewer and Burke (2002) found a weak effect of (mock) prosecution witness testimonial consistency on ratings of the likelihood of the accused committing the crime.

Other studies have also indicated that the types of language used could reveal deception, both in post hoc accounts of involvement in a mock crime event, and in email communication during the run-up to a pre-planned mock crime event. For example, liars have been found to become more self-focused, show greater negative affect, and evidence more cognitive processing compared to their truth-telling co-workers when in the process of committing insider crime (Taylor et al., 2013).

In aviation security settings, real-time identification of threat presents a huge challenge to those tasked with ensuring public safety, and to psychologists developing methods for detecting deception. Since the events of September 11, 2001, billions of dollars have been invested in aviation security procedures designed to detect threats to air travel, but the effectiveness of these procedures has been questioned (Weinberger, 2010). However, one new procedure based on testing the veracity of passengers' verbal accounts using techniques derived from much of the aforementioned psychological theory and laboratory studies has been found by the present chapter's third author to produce promising rates of deception detection, approaching 73%. This procedure, Controlled Cognitive Engagement (CCE; Ormerod & Dando, 2015), integrates theoretical and applied knowledge into a comprehensive procedure for detecting threat, which disregards behavioral indicators, concentrating on verbal indicators, instead. The results of a field trial conducted in several national and international airports revealed a clear advantage for veracity testing over the current method for aviation security screening. CCE detected more deceptive (mock) passengers without

increasing false alarm rates. Not only with mock but also genuine passengers, CCE interviews yielded more passenger talk and information. At the same time, security agents asked fewer but more effective questions. The use of CCE changed the verbal behavior of deceptive passengers, whose answers became shorter and had less information content by the end of the interview, while the answers of genuine passengers did not change. By using an information-gathering approach, first asking open questions about unpredictable topics that vary in their temporal reference, followed by test questions that seek information an individual should possess if they are being truthful, it is likely that CCE minimized cognitive demand for legitimate passengers but increased it for deceivers (Beckman, 2010). These findings provide evidence of the utility of verbal indicators, emanating from a structured, and managed interview technique.

Studies not of "Real Life"

A major criticism of almost all published studies involving attempts to detect truth and lies is that the stimuli (e.g., video recordings made available to assessors have not been of people lying in real-life, high-stakes situations; instead, studies often use students lying for the purposes of the experiment). Few studies have been able to examine in real-life legal/investigative settings which verbal cues might actually discriminate between truths and lies. In Sweden, Willen and Strömwall (2012) examined prisoners' accounts of (a) a crime they had actually committed (and been convicted for) and (b) a crime they had not committed but one they tried to convince the (research) interviewer that they had committed. Only the "clarity" of the accounts was found to discriminate.

In Norway, Myklebust and Bjorklund (2009) found that in court/legal proceedings regarding alleged child abuse a factor that discriminated between cases that resulted in convictions and cases that did not was the length of children's answers to open questions in police interviews. These findings suggest that detail was important. Further, a study in Italy (Roma, San Martini, Sabatello, Tatarelli, & Ferracuti, 2011) compared the contents of children's testimony in sexual abuse cases in which courts rendered either a guilty or not guilty verdict. The types of content that significantly differentiated between verdicts were quantity of details (including unusual and superfluous details), description of the interaction, reproduction of conversation and/or feelings, and spontaneous corrections/additions.

A Rare Study of Lying and Truth-Telling in Real-Life Police Interviews

Due to the ever-growing mutual respect between British police forces and criminal psychologists (that a number of psychologists have over the decades worked hard to achieve), we were able to secure comprehensive assistance

from a large police force in England to conduct a realistic lie detection study (Mann, Vrij, & Bull, 2004). This involved real-life police interviews with suspects that had been video recorded by the police and in which each suspect sometimes lied and sometimes told the truth. These recordings were observed for the purposes of our study by a large sample of police officers (not involved in the relevant investigations). We found an average lie/truth accuracy rate of 65%, with the lie detection rate being 66% and truth detection 64%. Furthermore, those officers who were more experienced in investigative interviewing performed better. Interestingly, officers whose beliefs about cues to deception fitted with lay person's stereotypical beliefs (e.g., gaze aversion, fidgeting) were the poorest at detecting lies/truths, but those who mentioned "content" cues (e.g., amount of detail, contradictions) were more accurate.

When we analyzed the behavior of the suspects in our study, we found lying to be associated not with the cues people commonly believe in, but with a decrease in blinking, an increase in speech pauses (i.e., that is pausing before answering a question or pausing during the giving of an account), and a decrease in hand/arm movements (for females) (Mann, Vrij, & Bull, 2002). Similarly, in their 2006 overview of studies examining which speech behaviors differentiate between truths and lies, Sporer and Schwandt found that among the very few cues found to discriminate were (i) response latency and (ii) speech errors.

The Crucial Importance of Skilled Interviewing

In his seminal 2003 publication, Kagan provided some advice on how interviews with asylum seekers should be conducted (e.g., to use rapport and open questions, and to avoid intimidation/confrontation). Given that judicial/legal evaluations may well be influenced by the ways in which witnesses, victims, and suspects give their accounts, and thus the ways in which they are questioned or interviewed, it would seem important that those who interview and question them are trained in ways which allow the accounts of witnesses, alleged victims, and suspects to contain factors known nowadays to indicate truthfulness. One example is the PEACE model of investigative interviewing that seeks to gain verbal content information from suspects, victims, and witnesses (see Milne & Bull, 1999 for a comprehensive review). It is important to note here that in their 2016 meta-analysis of training seeking to improve the detection of truths and lies Hauch, Sporer, Michael, and Meissner found that training based on verbal content cues was the most effective.

Time for a Change

Research studies conducted soon after the tape-recording of police interviews with suspects in England became mandatory in 1986 found mostly low levels of skills, which led the government and police chiefs to set up

(in 1991) a working party of highly experienced police investigators to develop up-to-date training on interviewing and interrogating. In light of their deliberations, which involved a focus on relevant psychology, they recommended what they called the PEACE model/approach. This change began in 1992 and it involved guidance documents and training courses that all police interviewers in England and Wales must attend, and which contained much research-based cognitive and social psychology (Milne & Bull, 1999).

While these detectives were having their working party meetings, the senior London police officer Tom Williamson convened a different small working party of detectives and psychologists (including Eric Shepherd, Stephen Moston, and the first author of the current chapter) that produced in 1991 an (unpublished) overview of aspects of psychology that might be useful to the improving of such interviewing/interrogating. This overview was made available to the national team of detectives that was developing PEACE. Once that team of detectives had drafted their guidance documents they sent drafts of these to the current chapter's first author asking if they had "got the psychology correct?" They indeed had.

Some of the (1992) Seven Principles of PEACE

The working party of experienced detectives drew up seven basic principles of the new PEACE approach that included:

> "The role is to obtain accurate and reliable information from suspects, witnesses or victims in order to discover the truth about matters under investigation." The focus is not on the mere gaining of confessions.

> "Interviews should be approached with an open mind. Information obtained from the person who is being interviewed should always be tested against what the interviewing officer already knows or what can reasonably be established." For our research on this disclosing of evidence/information by interviewers see below.

PEACE (see ACPO, 2013) is an acronym for:

P   Planning and Preparation
E   Engage and Explain
A   Account
C   Closure
E   Evaluation

Is the PEACE Model/Approach Effective?

In our study of 142 interviews with people suspected of fraud, we examined whether interviewing in a way that is compatible with the PEACE approach bore any relationship to the actual outcomes of these interviews

(Walsh & Bull, 2010). Overall, we found that better PEACE interviewing was associated with securing a greater number of comprehensive accounts, including exculpatory ones, as well as admissions and confessions. Also, see our other relevant studies (e.g., Walsh & Bull 2012a, 2012b), which examined the effects of rapport and how to overcome denials.

In 1992 the Home Office, part of the Government in England and Wales, published the pioneering research by Baldwin that it had commissioned to examine how interviews were conducted in the late 1980s after audio-tape recordings of them became mandatory in 1986. Of the 600 recorded interviews that Baldwin (1993) analyzed, "…most were short and surprisingly amiable discussions in which it often seemed that officers were rather tentative in putting allegations to a suspect…Indeed in almost two-thirds of all cases…no serious challenge was made by the interviewers to what the suspect was saying" (p. 331). Even when "the suspect denied the allegation, no challenge was made by the interviewers in almost 40 percent of cases" (p. 331). In only 20 of the 600 interviews that Baldwin examined did suspects "change their story in the course of an interview. In only nine of these cases was the change of heart attributable to the persuasive skills of the interviewer, and even here only three involved offences of any seriousness…The great majority of suspects stick to their starting position—whether admission, denial, or somewhere in between—regardless of how the interview is conducted" (p. 333). Pearse and Gudjonsson (1997) reported a similar finding.

Shepherd and Mortimer (1999) devoted a page of their chapter to the topic of deception and noted that while the "PEACE model of police interviewing requires" the interviewer to note "any unresolved inconsistency," this was "rarely done" at that time (p. 285).

When in Interviews is It Best to Confront/Challenge?

In her 2005 doctoral dissertation (that the first author of the current chapter supervised), Stavroula Soukara reported her study (see Bull & Soukara, 2010) of police tape-recorded interviews in which suspects actually shifted from denial to admission/confession. We wanted to examine the skills and tactics that the interviewers were demonstrating in the minutes that led up to the moment of these shifts. When in 2003, we first looked at her data we discovered that the interviewers were still "disclosing evidence," meaning that they had not given it all away at the beginning of the interviews (that used to be common practice in the UK, see Moston, Stephenson, & Williamson, 1992; and is still the case in some countries, see Lin & Shih, 2013). We then thought that the gradual or incremental disclosure of information and evidence could be important. Similarly, Hartwig, Granhag, Strömwall, and Vrij (2005) noted that research might be conducted on "different drip-feeding procedures in which parts of the evidence are disclosed throughout the interrogation" (p. 483). This is in line with one of the principles of the PEACE interviewing method mentioned above.

Some years later, having at last secured funds to support research studies on the possible effects of the gradual disclosure of information or evidence by interviewers, the first and third authors of the current chapter conducted a number of studies. In these studies (Dando & Bull, 2011; Dando, Bull, Ormerod, & Sandham, 2015), we found that gradual disclosure of information by interviewers led (i) themselves and (ii) observers of the video recorded interviews to be able to detect truth and lying at a rate of around 70%, this being greater than for early or late disclosure at 53% and 57%, respectively. We have also examined the significant effects of gradual disclosure on adolescents (Lingwood & Bull, 2013; McDougall & Bull, 2015) and in real-life interviews with people suspected of fraud (Walsh & Bull, 2015; also see Hartwig et al., 2011; Sorochinski et al., 2014). Such research sits well with what Meissner, Kelly, and Westerhof (2015) noted in their substantial overview on improving the effectiveness of suspect interrogations that "Research has consistently demonstrated that a suspect's perception of the evidence against them is an important factor that predicts true confessions" (p. 228).

A New Analysis

Most recently we have begun to examine which content cues might have been available within the interviews conducted for our 2015 publication (Dando et al., 2015) that enabled an unusually high truth and lie detection rate of around 70%. Here we present this examination for the first time. We decided to analyze the interviews for content cues that in DePaulo et al.'s (2003) meta-analysis of prior research on cues to deception had a considerable effect size, including *details*, *admitted lack of memory*, and *talking time*.

During our analysis for the *details* cue, it was noticed that interviewees sometimes gave reasons for their actions. Thus, it was decided to also include *reasons* in the analysis, which seemed to have not been studied in prior published studies. *Admitted lack of memory*, when an interviewee admits not to remember some specific details of an event, has been considered (within criteria-based content analysis—CBCA) to be an indicator of truthfulness (Vrij, 2005). However, interviewees sometimes fail to answer a question because they claim to have forgotten that information entirely. This type of admitted lack of memory seems to be different from the kind of memory loss that is defined within CBCA. In investigative interviews with suspects, different types of admitted lack of memory occur, from expressions that indicate uncertainty (e.g., "I might") to when an interviewee claims being unable to recall events completely. By solely looking at the CBCA admitted lack of memory criterion, one could ignore such other types of admitted lack of memory. Therefore, this study will look at several types of claimed memory loss and admittances of insecurity.

Method

The material used for the current analysis is from Dando et al. (2015). In that analysis, 150 graduate and postgraduate students were interviewed by an experienced police investigator who had undergone extensive specialist police interview training and also received additional training in the three information disclosure methods (i.e., early or late or gradual). The procedure for the participants consisted of two phases: first, they individually took part in an interactive computer simulation of building a major sports stadium for around an hour as either truth-tellers (i.e., mock builders) or deceivers (i.e., mock terrorists). Several aspects of each participant's computer activities (e.g., where they went to buy materials, which materials they bought, and the movements of their truck/lorry) were recorded on the computer and later accessed by a "covert" operative who shared some of this information with the interviewer. However, this information was far from enough to allow the interviewer to know if an interviewee was a builder or a terrorist.

The terrorists were to build but also later blow up the stadium. The two tasks were differentiated such that the terrorists had to try to mask their activities by trying to appear as legitimate builders. For example, while builders had to purchase small amounts of explosives to clear the site, terrorists were required to purchase much larger amounts, and thus had to find ways to hide their purchases. Terrorists were provided with an outline of the builders' tasks to allow them to devise ways of masking their true identity by appearing to be builders.

After the simulation had finished, the players were individually interviewed about their computer-based behavior using one of the three interview information disclosure methods. In each of these interviews, the interviewer used five of the pieces of information provided by the covert operative to disclose to the respective interviewee. These interviews were video-recorded and varying batches of 30 of these interviews were later shown to 30 lay observers (mean age 38 years, with no investigative or interviewing experience), whose task it was to decide whether each interviewee had been telling the truth or lying. The present study used 60 interviews randomly selected from this sample of 150. We included 20 interviews per disclosure method, of which half involved deceivers and half involved truth-tellers (i.e., a 2×3 design with ten interviews per cell).

Operationalization of the Dependent Variables

Talking time had two subcategories: *Talking time in the Free Recall Phase* (i.e., the duration of an interviewee talking during her or his free recall account) and *Talking time in the Questioning Phase* (i.e., the average interviewee talking time per question). Since this study involved calculating numbers of indicators per question, it was necessary to define a question.

A *question* was defined as being about one topic. All follow-ups, clarifications, and prompts are considered together as being part of that one question.

A *detail* was defined as a piece of investigation-relevant information provided that it related to "who, where, when, and what" issues. *Details* had two subcategories: *Details in the Free Recall Phase* and *Details in the Questioning Phase* (i.e., the average number of details provided per question).

A *reason* was defined as a piece of information that an interviewee used to give an explanation for his or her actions. This indicator also consisted of two subcategories: *Reasons in the Free Recall Phase* and *Reasons in the Questioning Phase* (i.e., average number of reasons per question).

Admitted lack of memory had four subcategories of (i) *I think* which is operationalized as the number of times an interviewee says "I think"; (ii) *Uncertainty* which was the number of times an interviewee admits he or she is not sure about his or her answer; (iii) *Fails to answer questions* which was defined as the number of times an interviewee is unable to answer a question; and (iv) a *Rest category* which included the number of other verbal indicators of uncertainty. For all of these four subcategories, the number is the total during the free recall phase and the questioning phase combined.

Results

In the questioning phase truth-tellers ($M = 2.37$) provided more details than deceivers ($M = 1.66$), this being in line with much prior research. Regarding saying "I think," deceivers ($M = 7.00$) said "I think" more often than did truth-tellers ($M = 4.00$), which aligns with the limited number of prior studies. Regarding the timing of disclosure (i.e., early or late or gradual), for *Details*, the number provided in the free recall phase significantly differed across disclosure method in that late disclosure ($M = 12.15$) was associated with significantly more details than gradual ($M = 7.35$); no difference was found between early and gradual. This substantial difference could be explained by those in the late disclosure condition not becoming aware of what the interviewer knew and thus not editing what they said.

Regarding saying "I think," the only significant effect was that the early ($M = 2.45$) and late ($M = 8.40$) conditions differed, perhaps because those in the late disclosure condition provided more details that "I think" could, in turn, be said about.

For *Reasons,* interviewees with whom the early disclosure was used provided significantly more reasons for their behavior ($M = 2.70$) than interviewees in the late ($M = 1.35$) and gradual ($M = 0.80$) conditions (no difference was found between late and gradual); this could be explained by those provided with early disclosure of information feeling that they had to give reasons for what they thus realized that the interviewer knew.

The admitted lack of memory, falling within the *Rest category,* revealed a significant main effect in that the early condition ($M = 0.75$) differed significantly from the late ($M = 2.00$) and the gradual ($M = 2.00$) (i.e., the late

and gradual did not differ from each other). A significant interaction was also found, in that truth-tellers ($M = 1.10$) and deceivers ($M = 2.90$) differed significantly within the late condition ($p = .01$), but not within the other two techniques. This could be related to the finding that in the questioning phase deceivers more often said "I think" than did truth-tellers.

As mentioned above, in our earlier study observers who saw the gradual disclosure interviews were better at detecting truths and lies. However, the new findings presented here that the gradual interviews did not contain significantly different numbers of verbal cues, from the early or the late disclosure interviews, indicate that the better truth/lie detection could not have been based on the observers using such verbal cues.

PEACE in Our Time?

In light of psychological research such that has been described above, a growing number of countries and organizations have decided to adopt the PEACE model/approach of investigative interviewing. Indeed, very recently (in summer 2016) the United Nation's (UN) "Special Rapporteur on torture and other cruel, inhuman or degrading treatment or punishment" (the Law Professor Juan E. Mendez) submitted his report that was then transmitted by the UN Secretary-General to the UN General Assembly. In this report, its summary stated that "The Special Rapporteur…advocates the development of a universal protocol identifying a set of standards for non-coercive interviewing methods and procedural safeguards that ought, as a matter of law and policy, to be applied at a minimum to all interviews by law enforcement officials, military and intelligence personnel and other bodies with investigative mandates."

When mentioning the "universal protocol" the UN Special Rapporteur noted that "Encouragingly, some States have moved away from accusatorial, manipulative and confession-driven interviewing models with a view to increasing accurate and reliable information and minimizing the risks of unreliable information and miscarriages of justice," and that "The essence of an alternative information-gathering model was first captured by the PEACE model of interviewing adopted in 1992 in England and Wales…investigative interviewing can provide positive guidance for the protocol…."

Furthermore, the Special Rapporteur stated that:

"…investigative interviewing…comprises a number of essential elements that are key to the prevention of mistreatment and coercion and help to guarantee effectiveness."

"Interviewers must, in particular, seek to obtain accurate and reliable information in the pursuit of truth; gather all available evidence pertinent to a case before beginning interviews; prepare and plan interviews based on that evidence; maintain a professional, fair and respectful attitude during questioning; establish and maintain a rapport with the interviewee; allow the interviewee to

give his or her free and uninterrupted account of the events; use open-ended questions and active listening; scrutinize the interviewee's account and analyse the information obtained against previously available information or evidence; and evaluate each interview with a view to learning and developing additional skills."

Since 1992, all of these skills have since been emphasized within the PEACE protocol/method we developed in England and Wales.

Concluding Comments

There are growing moves away from pressurizing and coercive interrogating that seeks confessions and toward interviewing in a manner designed to elicit comprehensive accounts, the contents of which can aid credibility judgments. This is especially important given that a meaningful proportion of those who have committed or are planning the most serious types of crimes are intelligent and socially skilled people, and it is such people who most often chose to lie and thus become skilled at doing so (Sarzynska et al., 2017).

References

Association of Chief Police Officers National Investigative Interviewing Strategic Steering Group. (2013). *Advice on the structure of visually recorded witness interviews* (2nd ed.). Available online at http://library.college.police.uk/docs/APPREF/ACPO-Witness-Interview-Structure-2013.pdf.

Baldwin, J. (1993). Police interview techniques. Establishing truth or proof? *British Journal of Criminology, 33,* 325–352.

Bell, B., & Loftus, E. (1988). Degree of detail of eyewitness testimony and mock juror judgments. *Journal of Applied Social Psychology, 14,* 1171–1192.

Beckman, J. F. (2010). Taming a beast of burden—On some issues with the conceptualization and operationalization of cognitive load. *Learning and Instruction, 20,* 250–264.

Bogaard, G., Meijer, E., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: Lay people's and police officers' beliefs about verbal and nonverbal cues to deception. *PLoS ONE, 11*(6), e0156615.

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26,* 353–364.

Bull, R., & Soukara, S. (2010). A set of studies of what really happens in police interviews with suspects. In G. D. Lassiter & C. Meissner (Eds.), *Interrogations and confessions.* Washington: American Psychological Association.

Campbell, B., Menaker, T., & King, W. (2015). The determination of victim credibility by adult and juvenile sexual assault investigators. *Journal of Criminal Justice, 43,* 29–39.

Canter, D., Ioannou, M., Youngs, D., & Chungh, G. (2016). Person perception aspects of judgments of truthfulness in public appeals. *Psychiatry, Psychology and Law, 23,* 547–562.

Dando, C., & Bull, R. (2011). Maximising opportunities to detect verbal deception: Training police officers to interview tactically. *Journal of Investigative Psychology and Offender Profiling, 8,* 189–202.

Dando, C., Bull, R., Ormerod, T., & Sandham, A. (2015). Helping to sort the liars from the truth-tellers: The gradual revelation of information during investigative interviews. *Legal and Criminological Psychology, 20,* 114–128.

DePaulo, B. M., Malone, B. E., Lindsay, J. J., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behaviour, 29,* 469–484.

Hartwig, M., Granhag, P. A., Strömwall, L. A., Wolf, A., Vrij, A., & Roos af Hjelmsäter, E. (2011). Detecting deception in suspects: Verbal cues as a function of interview strategy. *Psychology, Crime, & Law, 17,* 643–656.

Hauch, V., Sporer, S., Michael, S., & Meissner, C. (2016). Does Training Improve the Detection of Deception? A Meta-Analysis. *Communication Research, 43,* 283–343.

Home Office. (2015). *Asylum policy instruction: Assessing credibility and refugee status.* Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/397778/ASSESSING_CREDIBILITY_AND_REFUGEE_STATUS_V9_0.pdf.

Kagan, M. (2003). Is truth in the eye of the beholder? Objective credibility assessment in refugee status determination. *Georgetown Immigration Law Journal, 17,* 367–411.

Klaver, J., Lee, Z., & Hart, S. (2007). Psychopathy and nonverbal indicators of deception in offenders. *Law and Human Behavior, 21,* 337–351.

Lin, T. C., & Shih, C. H. (2013). *A study of police interrogation practice in Taiwan.* Paper presented at the Asian Conference of Criminal and Police Psychology, Singapore.

Lingwood, J., & Bull, R. (2013). Interviewing young adolescent suspects: When to reveal incriminating information? *The European Journal of Psychology Applied to Legal Context, 5,* 141–146.

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies, and videotape: An analysis of authentic high-stake liars. *Law and Human Behavior, 26,* 365–376.

Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology, 89,* 137–149.

Masip, J., Bethencourt, M., Lucas, G., Sanchez-San Segundo, M., & Herrero, C. (2012). Deception detection from written accounts. *Scandinavian Journal of Psychology, 53,* 103–111.

Masip, J., Blandon-Gitlin, I., Martinez, C., Herrero, C., & Ibabe, I. (2016). Strategic interviewing to detect deception: Cues to deception across repeated interviews. *Frontiers in Psychology, 7,* 1702.

McDougall, A., & Bull, R. (2015). Detecting truth in suspect interviews: The effect of use of evidence (early and gradual) and time delay on Criteria-Based Content Analysis, Reality Monitoring and inconsistency within suspect statements. *Psychology, Crime and Law, 21,* 514–530.

Meissner, C., Kelly, C., & Woestehoff, S. (2015). Improving the effectiveness of suspect interrogations. *Annual Review of Law and Social Science, 11,* 211–233.

Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice*. Chichester: Wiley.

Moston, S., Stephenson, G., & Williamson, T. (1992). The effects of case characteristics on suspect behaviour during police questioning. *British Journal of Criminology, 32,* 23–40.

Myklebust, T., & Bjorklund, R. (2009). The child verbal competence effect in court: A comparative study of field investigative interviews of children in child sexual abuse cases. *Journal of Investigative Psychology and Offender Profiling, 6,* 117–128.

Ormerod, T. C., & Dando, C. J. (2015). Finding a needle in a haystack: Veracity testing outperforms behaviour observation for aviation security screening. *Journal of Experimental Psychology: General, 144,* 76–84.

Pearse, J., & Gudjonsson, G. (1997). Police interviewing techniques at two south London police stations. *Psychology Crime and Law, 3,* 63–74.

Porter, S., & Yuille, J. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior, 20,* 443–458.

Riggio, R., Tucker, J., & Widaman, K. (1987). Verbal and nonverbal cues as mediators of deception ability. *Journal of Nonverbal Behavior, 11,* 126–145.

Roma, P., San Martini, P., Sabatello, U., Tatarelli, R., & Ferracuti, S. (2011). Validity of Criteria-Based Content Analysis (CBCA) at trial in free-narrative interviews. *Child Abuse and Neglect, 35,* 613–620.

Sarzyńska, J., Falkiewicz, M., Riegel, M., Babula, J., Margulies, D., Nęcka, E., et al. (2017). More intelligent extraverts are more likely to deceive. *PLoS ONE, 12,* e0176591.

Shepherd, E., & Mortimer, A. (1999). Identifying anomaly in evidential text. In A. Heaton-Armstrong, E. Shepherd, & D. Wolchover (Eds.), *Analysing witness testimony: A guide for legal practitioners and other professionals* (pp. 267–320). London: Blackstone.

Sooniste, T., Granhag, P., Stromwall, L., & Vrij, A. (2015). Statements about true and false intentions: Using the Cognitive Interview to magnify the differences. *Scandinavian Journal of Psychology, 56,* 371–378.

Sorochinski, M., Hartwig, M., Osborne, J., Wilkins, E., Marsh, J., Kazakov, D., et al. (2014). Interviewing to detect deception: When to disclose the evidence? *Journal of Police and Criminal Psychology, 29,* 87–94.

Sporer, S., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20,* 421–446.

Sweeney, J. (2009). Credibility, proof and refugee law. *International Journal of Refugee Law, 21,* 700–726.

Taylor, P., Dando, C., Ormerod, T., Ball, L., Jenkins, M., Sandham, A., et al. (2013). Detecting insider threats through language change. *Law and Human Behavior, 37,* 267–275.

Vrij, A. (2005). Criteria-based content analysis. A qualitative review of the first 37 studies. *Psychology, Public Policy and Law, 11,* 3–41.

Walsh, D., & Bull, R. (2010). What really is effective in interviews with suspects? A study comparing interview skills against interview outcomes. *Legal and Criminological Psychology, 15,* 305–321.

Walsh, D., & Bull, R. (2012a). Examining rapport in investigative interviews with suspects: Does its building and maintenance work? *Journal of Police and Criminal Psychology, 27,* 73–84.

Walsh, D., & Bull, R. (2012b). How do interviewers attempt to overcome suspects' denials? *Psychiatry, Psychology and Law, 19,* 151–168.

Walsh, D., & Bull, R. (2015). The association between interview skills, questioning and evidence disclosure strategies, and interview outcomes. *Psychology, Crime and Law, 21,* 661–680.

Weinberger, S. (2010). Intent to deceive? *Nature, 465,* 412–415.

Willen, R., & Stromwall, L. (2012). Offenders' lies and truths: An evaluation of the Supreme Court of Sweden's criteria for credibility assessment. *Psychology, Crime and Law, 18,* 745–758.

Disbelief Repeats as Deception Tagging: Conversational Strategies for Labeling Perceived Deception in Interrogation

Gary C. David and James Trainum

When suspects lie in police interrogations, it is the job of investigators to identify those lies. However, increased concerns around confrontational interrogation tactics means that investigators also need to be mindful of anything that could be seen as coercive. In this chapter, we explore subtle attempts to tag deception, namely through the conversational strategy of disbelief repeats. Using conversation analysis on actual interrogations, this chapter identifies how *disbelief repeats* shape suspect responses through a less aggressive, but impactful approach. Ultimately, this chapter emphasizes the need to look at suspect deception not simply as a psychological state, but as part of a communicative interaction between suspects and investigators.

Introduction

For investigators, an important element of their role in an interrogation is to identify when suspects are being deceptive. As Stokoe (2010) states, "Unsurprisingly, the suspects' main business in police interrogations is often denying

G. C. David (✉)
Bentley University, Waltham, MA, USA
e-mail: gdavid@bentley.edu

J. Trainum
Criminal Case Review and Consulting, Washington, DC, USA

the charges put to them" (p. 63). Suspects' denials are often complex and go far beyond saying "I didn't do it." They extend to giving answers and alibis that provide incomplete information, shadow the truth, and outright lies (Moston & Stephenson, 2009; Olson & Wells, 2004). Knowing that suspects typically want to avoid being found guilty, there can be little reason to trust what a suspect says when denying involvement in a crime. When conducting interrogations, investigators are not just passive recipients of information provided by suspects; they are active participants, jointly constructing the interaction with the suspect. As interrogations are a type of turn-taking system (Sacks, Schegloff, & Jefferson, 1974), the acts of investigators and suspects produce the final product of what is recognizable as an interrogation. Part of this process includes how investigators handle suspects who are perceived to be lying.

In terms of detecting lies, Carter (2014) notes, "Lies in police interviews remain under-explored from a conversation analytic perspective…" (p. 124). We extend this perspective to explore how investigators react when they believe suspects are lying. In everyday life, there are very limited means of detecting deception and ultimately depend on factors such as power relations, context of the encounter, and the significance of the lie. When individuals encounter what they perceive as deception, their responses may range from ignoring it to confronting their interaction partner directly. Direct confrontation carries with it an implied commitment to deal with the repercussions of the accusation. Given that "everyone has to lie" (Sacks, 1975) about even the most rudimentary things, directly addressing every lie can be burdensome. The same can be said about police reactions in interrogations. Police can decide to overlook parts of a story that they believe to be not true if the lies are not important to the case. Conversely, police can directly challenge a suspect whose claims diverge from what is believed to have happened.

This chapter explores *subtle tagging* of deception, in which investigators do not directly challenge a perceived lie by labeling it as such. We refer to this conversation-based technique as a *disbelief repeat*. Repeats are a common feature of conversation, commonly arising when some issue arises between what was said and what was heard. However, unlike other types of repeats, the disbelief repeat functions to mark a suspect's response as problematic. Interrogations, like all conversations, involve a turn-taking process where one person's utterance leads to a number of follow-up responses (Sacks, Schegloff, & Jefferson, 1974). As such, a direct and explicit accusation of dishonesty *can* impact what happens next in the encounter, as can more subtle means of delivering the same message. In this chapter, we focus on how investigators use disbelief repeats, and how these conversational moves can shape interrogations.

LABELING DECEPTION

Simply put, deception can be defined as "a deliberate attempt to convince someone of something the liar believes is untrue" (Granhag, Vrij, & Verschuere, 2015, p. xv). In terms of deception as a strategy of conflict—which we consider interrogation often to be—deception is when information is designed "to manipulate the behavior of others by inducing them to accept a false or distorted presentation…" (Whaley, 1969, p. 188; see also Gerwehr & Glenn, 2000). Ultimately, "a lie isn't simply any old falsehood; it's told with the knowledge that it's false and with the intent to deceive" (Greenberg, 2017). An extensive amount of research has been devoted to determining when a suspect is being deceptive (DePaulo et al., 2003; Granhag et al., 2015; Vrij, 2000). At the same time, disagreement still exists regarding how useful and reliable any of these approaches are to identifying deception (Bond & DePaulo, 2006; Dando & Bull, 2011; Vrij, 2004; Vrij, Meissner, & Kassin, 2015).

Saying someone is being deceptive is difficult since you would have to know whether or not that person is intentionally lying, which ultimately would involve getting into that person's mind and intent. Accusing someone of lying, therefore, becomes delicate business. Making such an allegation carries with it a tremendous weight, charging the person who made the questionable statement as having a dubious moral character. In many contexts, there exists a hesitancy to explicitly call someone a "liar," even if evidence proves the deception.

US politics, for example, present a recent and useful example of this hesitancy to label someone a "liar." Individuals generally expect that all politicians will lie or "misstate the facts" to some extent. The Pew Research Center poll on "Public Trust in Government" shows historical lows, with only about "18% of Americans today saying they can trust the government in Washington to do what is right 'just about always' (3%) or 'most of the time' (15%)" (Public Trust in Government, 2017, para. 1). Political fact-checker Angie Drobnic Holan (2015) in the title of her news article states, "All politicians lie. Some more than others." For instance, Kessler, Rizzo, and Kelly (2018, para. 1) claim, based on an analysis of the fact checker database, that "President Trump has made 3001 false or misleading claims," which averaged to around 6.5 claims a day. In one instance, President Trump claimed in a November 21, 2015, speech that "thousands and thousands of people were cheering" in Jersey City, New Jersey celebrating the attacks on September 11, 2001. Despite no video evidence, no other eye-witness accounts, no police reports, and no other evidence of any kind, when challenged by ABC *This Week* host George Stephanopoulos, then-candidate Trump restated the assertion, claiming he saw it on television. On November 23, he tweeted out

a link to a news story that he claimed to support this assertion, but in fact did not.

Despite this pattern, the media, overall, has appeared to struggle in terms of how to label those whom deceive. The media generally has shown hesitancy to label a politician as an outright liar (Beinhart, 2017). Rather than making direct assertions that someone is lying, alternative labels have been used, such as "falsehood," "untruth," and "alternative facts" (Waldman, 2017). Arthur Brisbane, Public Editor of the *New York Times*, even went as far as to ask readers "whether and when *New York Times* news reporters should challenge 'facts' that are asserted by newsmakers they write about" (2012, para. 1). Fallows (2016) calls for aggressively challenging lies told out of the White House in order to demonstrate what is known to be objectively true as a way of establishing a "shared reality."

The issue of a "shared reality" refers to a fundamental goal in interaction, which is the achievement of intersubjectivity (or shared meaning). From an ethnomethodological perspective, shared meaning is accomplished in and through interaction (Garfinkel, 1967, 1948/2006). Based on this perspective, deception is not just a matter of psychological states, but interactional ones, as well. In other words, focusing only on "lie telling" is just half of the equation. Equally important is the reaction of those who believe they are being lied to. In everyday dyadic situations, direct questioning of a perceived lie might not be a practical option. However, in institutional encounters, where power is unequally distributed, directly addressing deception becomes a more possible option, as there is an uneven power differential between interlocutors. For example, a patient can make false claims that he or she followed a restricted diet, but a physician can claim noncompliance through evidence to the contrary. In another example, a student can claim that their dog ate their homework, but the teacher can choose not to accept this excuse. In each situation, the receiver has the authority to express that the sender is engaging in actual or potential deception. What versions of reality are privileged as being true can be based on the power that one possesses in that setting. Thus, the goal of institutional interactions is not necessarily to establish a shared reality but a *preferred* reality, and one that privileges the points of view of those with institutional power. Generally, for law enforcement interrogators, what constitutes a preferred reality is often a version of events that places the suspect as guilty. Any information that detracts from that version of events may end up as being treated as deceptive.

Research in the field of conversation analysis has examined how lies are told, and how receivers react (Carter, 2014; Edwards, 2006; Komter, 2003; Stokoe, 2010). Using transcripts primarily from reality-style television shows, Reynolds and Rendle-Short (2011) (see also Reynolds, 2009) closely examined how the actions of senders and receivers collaboratively built the telling of lies, as well as the revealing of one's dishonesty. The shows included *COPS* (the original reality show meant to portray what policing is like from

the standpoint of the police) and *The Jeremy Kyle Show* (a UK television talk show which involves two or more parties in a dispute). In these shows, perceived and actual lies were treated in one of three ways. First, the person who told the lie could "come clean" and admit that he or she had in fact lied. In this instance, some kind of evidence (e.g., a polygraph or DNA test) could be presented that does the work of showing the person was lying. Second, other participants in the encounter labeled what was said as being a lie. The person accused may continue to deny that deception is taking place and stick to the original story.

This led to the third way in which deception proceeded: The person who told the lie changed his or her story based on how others reacted to it. Here, "lying is made relevant by the participants and the lie-teller subsequently revises their response under further re-questioning" (2009, p. 62). The fact that others are challenging the lie-teller's statements results in revisions by the person telling the lies. This last element relates to our focus here: How reformulating statements as questions results in tagging these statements as deceptive and calls the speaker to reconsider his or her response.

Repairs, Corrections, and Disbelief Repeats

Conversation analysts note that *repairs* and *corrections* are common elements of everyday conversations, occurring when some type of breakdown or misunderstanding happens in an interaction. When a breakdown happens, it is common for those involved to try to "repair" the conversation (Schegloff, 1991, p. 157). This can be a collaborative activity where those involved work toward the common goal of setting things on a proper track. A *correction*, however, "is commonly understood to refer to the replacement of an 'error' or 'mistake' by what is 'correct'" (Schegloff, Jefferson, & Sacks, 1977, p. 363). A repair can involve subtle work where minimal attention is brought to the breakdown. A correction occurs when someone issues a direct statement to someone saying or doing something "wrong."

In terms of their conversational positioning, corrections and repairs occur in a variety of places. Heritage (2013) remarks, "Within conversation analysis it is conventional to distinguish between first (initiating), second (responsive), and third (sequence closing) positions in a sequence" (p. 331). When trying to repair an interactional problem in the *first position*, a speaker can immediately correct a mistake. In fact, in conversation there is a "preference for self-correction" (Schegloff et al., 1977), allowing the speaker to fix an error with minimal prompting. In the *second position*, Robinson and Kevoe-Feldman (2010) remark that such a response can "help speakers, who will be engaged in the process of repair, answer two questions: (a) what needs to be fixed?, and (b) how should it be fixed?" (p. 232). Finally, a *third-position repair* can be performed by the original speaker, who corrects the original problem, thereby closing out the repair sequence.

In this chapter, we focus on corrections that occur as responses in the second position, what we call *disbelief repeats*. People immediately repeating what was just said are a regular feature of talk (Drew, 1992a, b; Jefferson, 1972; Robinson, 2012; Robinson & Kevoe-Feldman, 2010; Schegloff, 1997; Schegloff et al., 1977). Svennevig (2008) identified three general types of repeats: (1) hearing, (2) understanding, and (3) acceptability. Problems of hearing occur when the repeat serves to indicate that something may have not been heard correctly, and thus needs to be repeated. For instance, a wife may tell her partner that she will be home from work at 6:00 p.m. Due to the commotion caused by children playing loudly in the background, her partner may repeat "You'll be home at 6:00 p.m.?" The repeat simply serves to confirm a correct hearing.

Problems of understanding, on the other hand, refer to instances where a hearer may not understand what is meant by what was said, functioning as a display of incomprehension. In the same scenario, the partner may similarly ask "You'll be home at 6:00 p.m.?", due to the fact that the wife typically works till 9:00 p.m. on that night. The wife might then say, "Yes, because my later meetings were cancelled this week," thus giving the basis for the change in routine. Finally, problems of acceptability indicate that the speaker may have made an error and thus needs to correct what was said. Once again, upon being asked "You'll be home at 6:00 p.m.?", the wife may self-correct by saying "Oh I mean 9:00 p.m. I forgot today is my late night."

Thus, the same repeat of "You'll be home at 6:00 p.m.?" can serve different functions. Interestingly, in Svennevig's analysis, he found that even when problems of acceptability were evident, people responded by positing it as a problem of hearing. In other words, mistakes made by the speaker were acted upon as if the person misheard or misunderstood what was said (even though that was not in fact the case). Or, in the words of Pomerantz (1984), speakers "try the least complicated and costly remedy first" (p. 156). Rather than correcting the wife that she had the wrong day, the repeat "You'll be home at 6:00 p.m." to allow the speaker to save face by initiating a self-repair, rather than being corrected as having the wrong day.

Part of the reason for this attempt to allow self-repair versus correction relates to the concept of *facework* (Goffman, 1967). Essentially, facework refers to a presentation of self that provides a positive image of the person's internal state and external expression of that state. The idea of "saving face" refers to allowing one to do the work of correcting a problem on his or her own, which indicates the person recognizes their own error. Thus, when a conversational problem occurs that is in need of repair, the person listening presents an opportunity for the speaker to do the repair work (or take corrective action) in a way that can minimize problems, lessen embarrassment, and allow for the interaction to proceed. Repeats can preserve the delicate work of interaction, providing a space for the speaker to adjust what

was said in such a way that the conversation can move forward on the basis of a shared understanding of events.

Regarding police interrogations, repeats can come in the form of repairs that fit the previously discussed classifications. Additionally, we argue that repeats can occur in the form of corrections, seeking to label a response as erroneous, misleading, or deceptive. Through the disbelief repeat, the investigator is able to indicate doubts about what is being said in a way that does not aggressively ask for a correction, but allows the speaker to provide a repair. For instance, a suspect might give a response that the police do not believe to be true (or know to be false). The police might say, "That's not what happened. Now tell the truth." This instructs the suspect to issue a correction. On the other hand, by repeating what was said in the form of a question, the police can provide the suspect with a chance to repair his or her statement. Since repeats are not seen as direct accusations of lying, issuing a repeat can be a kinder, gentler way of requesting a correction.

The dynamics in a police interrogation are not those we typically find in other parts of everyday life. The differential power and aggressive tactics, especially in the context of US policing, can make the repeat function precisely as a correction, possibly acting as a type of *camouflaged coercion*. Because the police are showing their rejection of what is being said, this response can influence what the suspect says next. Leo and Drizin (2010) define interrogation as a two-step process. First, through the use of interrogation techniques meant to persuade the suspect to confess and convince him or her of guilt, the suspect is led to believe that he or she is caught and powerless to change the situation. Second, suspects are convinced that compliance is in their best interest. The disbelief repeat can be used as part of this process by correcting what the suspect says in what appears to be non-confrontational terms. It also should be remembered that police are in the business of assessing moral character (Sacks, 1972). The maintenance of "face" for the suspect becomes less of a consideration, if a concern at all. Police then may not be issuing a repeat to maintain the delicate work of interaction, but rather to get the suspect to change his or her story. Such a goal makes what appears to be a "polite" repair into a potentially coercive element of police interrogation.

Tagging deception in the form of a disbelief repeat thus is an important part of the interrogation process. Carter's (2011, 2014) extensive research into police interrogations has explored "the interactional manifestation of lies and deceptive interaction…focusing on both the lie and the subsequent responses produced by the suspect when the lie is explored by the police" (2014, p. 123). Explicit claims of deception have received the most attention from researchers and lawyers. Investigators can (and do) come straight out and claim a suspect is lying. At the same time, doing so carries a number of potential risks. First, there is the potential of losing the suspect's cooperation in continuing the interrogation, especially if no lawyer is present. Second,

these types of accusations can damage the process of rapport building and weaken the establishment of trust with the suspect. Third, aggressively issuing claims of lying can potentially create the appearance of coercion, especially if the interrogation is being recorded. Contrary to this approach is what we call *subtle tagging*, in which investigators mark statements as being problematic without direct accusation of lying. Despite its appearance of being more "gentle," it nonetheless impacts how interactions proceed, and how subsequent suspect statements are shaped.

Methods and Data

Our approach aligns with Vrij's (2000) call to explore deception beyond experimental designs in laboratory settings and to examine the nuances of real-life deception. The primary data utilized for this study included eighteen videotaped interrogations, which typically included official transcription of the interrogation. Eighteen of the interrogations were examined as part of the authors' work as case consultants for defense counsel. As part of this work, the authors were asked to perform analysis of the interrogations (and in some instances police reports and other related case materials) to determine the extent to which coercion, contamination, and/or false confessions were present. Video or audio recordings, along with a transcript, were provided as part of this work. Supplemental materials, such as police reports, statements, medical examiner findings, and the like, were also utilized when available. Since this work was initially performed as consulting and not research, no IRB approval was needed to perform this work. However, we secured permission as part our routine agreements to use anonymized elements from the interrogations for training and educational purposes. Additionally, we examined 14 cases gathered from publically available sources, such as Garrett's (2011) case database (www.convictingtheinnocent.com) and the National Registry of Exonerations. These cases all involved felony crimes, such as murder, robbery, rape, and child molestation. All identifiers have been removed to preserve anonymity of those involved.

A conversation analytic-based approach was used to examine the videotapes and transcripts, in order to focus on the *practices*, *sequences*, and *systematic organization* of the interactions (Drew & Heritage, 2006). Such an approach is rooted in applied conversation analysis (Antaki, 2011) and the examination of institutional talk (Drew, 1992a, b). Our primary focus was on broader conversational features within the interaction, as opposed to very detailed examinations of microelements that are more prevalent in conversation analysis. Specifically, we were interested in understanding how police responded when they believed the suspects were being deceptive. Utilizing Carter's (2014) approach, our attention was "directed towards the influence of the questioner's talk on the deceiver's response (in order to) provide a more useful understanding of the manifestation of deception…" (p. 137).

The interrogations were all conducted in the US. There is no one mandated way in which interrogations are done, especially given the decentralized policing structure in the US. While there are general legal guidelines regarding how an interrogation should, and should not, be performed, what this means in practice can be highly variable. Thus, it is difficult to determine how representative any one, or any set, of interrogations is of interrogations in general. However, they do reflect how interrogations can be and are done. While no single type of interrogation exists, the dominant approach utilized is the Reid technique, taught in various forms by most interrogation schools. The Reid technique involves a behavioral analysis interview, as well as a 9-step interrogation process (Inbau, Reid, Buckley, & Jayne, 2005). The use of deception is a recommended technique in the approach, where investigators make claims of evidence that may not exist and knowledge of the crime that they may not actually have. The Reid Institute refers to this as "misrepresenting evidence" (Reid Institute, n.d.). Also, the Reid technique advises interrogators on how to deal with suspect denials, specifically by repeatedly encouraging the suspect not to lie and tell the truth (Inbau et al., 2005). While we were not able to confirm in every instance that the investigators were Reid trained, features of the Reid technique were present in most of the interrogations we examined.

DISBELIEF REPEATS AS TAGGING DECEPTION

A suspect, allegedly having primary knowledge of a crime, is assumed to know more than the police, thereby creating an "epistemic gradient" (Heritage, 2012) where one party has more knowledge than the other. The investigator, not having been "at the scene of the crime," must rely on what Pomerantz (1980) refers to as a Type 2 knowable or something known by report, hearsay, or inference. Investigators, however, are often not at a complete disadvantage in terms of their knowledge. Forensic data and other evidence can help to lessen the gradient (Heritage, 2012), allowing a window into events independent of what the suspect says. Interrogation can then become akin to a negotiation over who knows what, and whose version of events is more believable. Being seen as believable is an interactional accomplishment. Part of this effort lies in how people come off as believable through the interaction. If a person is in some way stigmatized, he or she may have a very difficult task in being believed. Even when being truthful, such a person can still be seen as being deceptive.

The content of the story being told also has an impact on whether a person is seen as believable. In the case of an interrogation, the investigators often provide their own perspective on the suspect's story as it is being told. Investigators can play the role of *assistant storyteller*, helping to form the suspect's story based on what they know to be forensically true, what they believe to be true professionally, and what the suspect will reveal truthfully.

Investigators generally believe that in order to get a truthful narrative, they must rely on training to get at the ground truth, which suspects can be resistant to share.

Some interrogators are trained to aggressively, verbally engage suspects who they believe are lying (e.g., some elements of the Reid technique) or take a more circuitous route of asking repeated follow-up questions to explore inconsistencies or problems in a suspect's answer (e.g., as more often done in the PEACE [Preparation and Planning, Engage and Explain, Account, Closure, and Evaluate]) Method practiced in the UK and Canada. On the verbally aggressive end of the spectrum, suspect statements that are disbelieved can be met with exclamations of "Stop lying to me!", "That doesn't make sense!", or "That's not what happened!" Accompanying this can be references to real or imagined forensic evidence to bolster the rebuttal provided by investigators. Furthermore, investigators can make moral judgments of lying, which can reflect poorly on the suspect.

The example below demonstrates how an investigator can directly accuse a suspect of lying and looking "bad." In this instance, taken from the authors' own dataset, the suspect (whose name has been changed) is being accused of murdering his boyfriend. The suspect has maintained his innocence by claiming the victim jumped from the balcony after a mental health episode. The police believe that the suspect threw him from the balcony after a violent confrontation. By telling the investigator a version that matches what is believed to be true, the suspect can redeem himself in the eyes of the investigator.

> **Extract #1** (Case 7; p. 56 of interrogation transcript)
> Steve, look at me. I don't want to hear any more lies. Okay. Because you have been lying okay? This is a serious situation, okay? You cannot back talk your way out of it. Okay? You need to be 100 percent honest with us and let us help you through this situation. This is what we do. I don't think you're a bad person. I don't think that whatever happened in the apartment was your intent. I think things got out of hand and out of control. But if you continue to lie, people will see you as being evil and meaning to do something. Okay. Lying is the wrong thing to do in this situation, okay, because it makes you look bad and you don't want to make yourself look bad right now.

At the beginning of the extract, the investigator directly goes at Steve's denials of guilt. Interrogation manuals suggest taking an approach that directly confronts and attempts to overcome denials. Gordon and Fleisher (2011) recommend that investigators "do not allow the suspect to deny the act" (p. 258). Inbau et al. (2005) suggest that the investigator should ignore weak denials, but when a strong denial is issued "should reassert his confidence in the suspect's guilt" (p. 166). Of course, if an investigator believes that a suspect is guilty, anything said that does not confirm that guilt is going to be seen as deceptive.

Even though investigators may want to directly challenge a suspect and accuse him or her of lying, they also need to be careful not to attack the suspect. This is especially true in an era of increasing video recording of interrogations. Rabon and Chapman (2009) recommend to "[*n*]*ever* personalize the interview process" (p. 12; emphasis in original). Doing so can lead to reacting emotionally and thus losing control of the interrogation. Hess (2010) recommends incorporating protests that dispute evidence into a larger narrative, providing the suspect "with acceptable reasons to confess" (p. 79). This is similar to the idea of *theme development* in the Reid technique, in which the investigator provides a rationale for the suspect, which reduces the perceived severity of the crime and presents the suspect in the best light possible. By limiting overly antagonistic responses, the suspect will be more likely to buy what the investigator is selling: the idea that confessing is good for the suspect.

Disbelief repeats function as a way of demonstrating that what was said is questionable in terms of truthfulness, but without directly attacking the suspect. By reframing the response in the form of a question, the suspect can be alerted to the suspicion of deception. What also is important about this type of formulation is that it appears to be relatively innocuous. The structure of the disbelief repeats look like any kind of conversational repeats and might appear to act in the same way. For instance, it might appear that the investigator does not hear what the suspect said or is unclear on what was said. Moving from the form and into function, however, it becomes clear that the repeats act in a different way. It effectively tags the utterance as needing correction since it is being perceived as being deceptive.

In the extract below, taken from a case that the authors worked on, the investigator is questioning a suspect who is believed to have been indirectly involved in a murder and robbery. She is alleged to have picked up her boyfriend Ernie after he is suspected of robbing and then murdering a victim. The investigator wants to know when she confronted her boyfriend Ernie about the murder:

Extract #2 (Case 8; p. 13 of interrogation transcript)

Investigator:	And at what point did you confront Ernie?
Suspect:	I never said nothing to Ernie
Investigator:	**You never said nothing to Ernie?**
Suspect:	(Unintelligible.) Like, when I looked at the area, I was like, to myself, the area right where I was at, but –
Investigator:	You see, I – here's where I have a hard time believing you there…

The suspect responds that she "never said nothing to Ernie." Immediately, the investigator uses the exact formulation of the suspect's statement, but reframes it as a question. This prompts the suspect to expand on her original statement with an account. As she does so, the investigator cuts her off,

stating his disbelief. This example demonstrates how the repeated statement is meant to tag the original statement as deceptive. The investigator characterizes his mental state as one of disbelief (e.g., "I have a hard time believing you there…"). The "there" is sequentially related to the disbelief repeat, which tags the original statement as the point of contention.

Further, the below example from another case of the authors demonstrates how disbelief repeats are performed as rephrasing the answer as a question. The suspect here is accused of being part of a group paid to collect money from a person. If the person refused to pay, they were then to kill the person who owed the money. The investigator is asking if the suspect knew about the origins of the money that was owed:

Extract #3 (Case 9; p. 16 of interrogation transcript)

Investigator:	…that this money…did he tell you why he had to get it? Was it for business, was it for drugs, was it because he had….?
Suspect:	He didn't tell me. I didn't ask him. He just told me that he was going to give me 2 thousand bucks, or something like that.
Investigator:	**And you didn't know why he had to collect the money?**
Suspect:	No, see, I just did him the favor of bringing him in my car to give him a ride 'cause he didn't have a car.

The suspect answers the question that he never inquired regarding the origins of the money that was owed. Rather, he was just going to get paid about two thousand dollars. The investigator, not satisfied with the answer, continues to probe, forming the suspect's answer into a question (e.g., "And you didn't know why he had to collect the money?"). Even though the suspect stated, "He didn't tell me. I didn't ask him," in the *previous* conversational turn, the investigator's disbelief repeat culls additional information from the suspect. The suspect's response to the investigator's disbelief repeat is noteworthy, as the suspect expands the detail of his account, helping construct a potentially more telling, truthful account of events.

In an interrogation regarding the same case with the same suspect a few moments later, the pattern of question-response-repeat occurs again:

Extract #4 (Case 9; p. 16 of interrogation transcript)

Investigator:	Did you know…who he was coming to collect the money?
Suspect:	He didn't tell me.
Investigator:	**He never told you?**
Suspect:	No, he didn't tell me.

The suspect again demonstrates limited knowledge regarding the crime, saying that he was not told from whom they were to collect the money. This leads to the investigator to re-query the response, taking the response and forming it into a question, "He never told you?" It is noteworthy that after

the suspect states, "He didn't tell me," the investigator *elevates* his response to "He *never* told you?" The investigator's disbelief repeat indicates a kind of permanence of position regarding the suspect's lack of knowledge. To this re-query, the suspect provides no expanded explanation.

Disbelief repeats can function as a way to gently challenge suspects without being aggressive, accusatory, or threatening. Part of this approach is owed to the ambiguity of the repeat. Is the person repeating the statement as a challenge, to seek confirmation of what was said or to seek clarification? How the suspect interprets the disbelief statement can be reflected in the turn-taking of suspect-investigator talk, and thus the impact of the disbelief repeat can also be evaluated from how the interactants respond to the repeat. In the below example taken from our dataset (case 7, which was also utilized for Extract #1), the suspect is accused of murdering his boyfriend "Dennis" by pushing him off of an apartment balcony. The following exchange takes place at the second time the suspect was questioned about the death of his boyfriend. At this point in their dialogue, the suspect has raised that the victim was acting erratically on the night of his death and that it was he who was afraid of being a victim.

Extract #5 (Case 7, p. 54 of interrogation transcript)

Investigator:	Did he come at you?
Suspect:	He did not try to get me, but all that I did was I tried to physically wake him up by just - - I shook him first - -
Investigator:	I know. You - -
Suspect:	And he didn't respond to me.
Investigator:	You told us that. We're past that. Okay? Did he come after you? You said you were afraid of him and he was angry. What made you afraid of him? Why were you scared of him?
Suspect:	Because he - - the way that he was screaming, I thought he was going to come after me. And, and - -
Investigator:	Was he mad at you?
Suspect:	Not that I know of. He never got mad at me.
Investigator:	**He never - - he's never gotten mad at you?**
Suspect:	Except for just four things. And I can't - -
Investigator:	You don't need to remember them right now. Dennis told a lot of his friends that you were aggressive.
Suspect:	**Aggressive?**
Investigator:	Aggressive. Did he ever tell you you were aggressive?
Suspect:	No, never told me I was aggressive…

At the outset of this exchange, it would appear that the investigator is entertaining the suspect's claim that he was afraid of the victim's erratic behavior. Quickly, though, it becomes clear that the suspect's version of events is not being accepted. The investigator cuts off the suspect's response, which was answering the question "Did he come at you?" This is an important question, as it is a scenario constructed by the investigator for the suspect

to then address. A few moments earlier, the suspect offered that he was afraid of how the victim was acting (e.g., "I was afraid he was going to hurt me because of the way he was - -"). The investigator interrupts and introduces the question of whether the victim tried "to hurt you and did you have to defend yourself."

The suspect responds that "He did not try to get me…" The only contact he states they had was when "I tried to physically wake him up by just - - I shook him at first - -." The investigator cuts off the suspect's response, indicating that this has already been discussed, and that "We're past that." Through this conversational move, the investigator is able to maintain control over the topic being discussed. The question "Did he come after you?" is restated, with an additional account: "You said you were afraid of him and he was angry." This places the suspect's words at the locus of the question, indicating the relevance of the question based on his earlier utterance.

Again, the suspect begins to answer the question asked by the investigator. He indicates that the victim was screaming and that he "thought that he was going to come after me." The suspect does not indicate that the victim in fact came after him, but that he *thought* he was going to do so. Clearly, this is an important distinction, because the investigator is constructing a scenario in which the suspect indicates self-defense due to the aggression of the victim. As the suspect starts to expand his response, he is again cut off by the investigator, who asks "Was he mad at you?"

The question explores whether some issue existed between the suspect and the victim, which might have precipitated conflict and violence. The suspect indicates that he was afraid of the victim due to how he was acting at that moment, which was said to be not typical behavior. The investigator builds a scenario in which the conflict was more commonplace. These two versions of events come to a head shortly. To the question of whether the victim was mad, the suspect responds, "He never got mad at me." Stating that the victim never got mad at the suspect contradicts the investigator's belief that the couple was fighting as a result of something that happened.

Rather than directly challenging the suspect for lying, the investigator restates the suspect's statement as a question: "He never - - he's never gotten mad at you?" The suspect appears to perceive the question as a challenge to the veracity of his earlier statement (e.g., "He never got mad at me"), as the suspect begins to qualify the original statement to identify "just four things" that made the victim mad at him previously.

 Once again, the investigator cuts off the suspect, saying that he does not need to recall the "four things" at this moment. His lack of interest might be considered strange given he was seeking the foundation for why the victim might have been mad at the suspect and aggressive toward the suspect. However, the reason for this becomes clear in the next statement: "Dennis told a lot of his friends that you were aggressive." Here, the investigator changes locus of aggression from the victim to the suspect. This constitutes a type of

news delivery (Maynard, 2003), given without any warning of the topic being shifted from the victim's behavior to the suspect. The utterance was not asked as a question, but rather as a statement of fact as reported by "a lot of his friends." Here, we see the ability of investigators to speak in general or non-specific terms, something that suspects often do not have the ability to do since they are held to what they say to a higher standard than are investigators.

In response, the suspect repeats the word "Aggressive?" It is difficult to say whether the suspect is indicating his disbelief through this repeat, or perhaps confirmation of what was said (or both). The investigator repeats the word "Aggressive," followed up with the yes or no query, "Did he ever tell you you were aggressive?" The suspect, adhering to the structure of the question, responds, "No," and then copies the formulation of the investigator, indicating, "…never told me I was aggressive…" Despite the investigator's attempts to reformulate the suspect as the aggressor, the suspect resists this characterization. In the end, however, the suspect does confess to the crime and did go to prison, even though he still contests his confession as involuntary and coerced.

These examples provided in the extracts demonstrate the range of disbelief repeats as they function in actual interrogations. They come in conjunction with what Antaki (1994) refers to as *explanations in exoneration*, which are given "when we are in some trouble, or faced with some kind of accusation" (p. 43). When a suspect is being questioned regarding involvement in a crime, the suspect must be perceived as being truthful in order for responses to be accepted. Likewise, investigators are engaged in a game of trying to ferret out deceptive responses, which are typically defined as responses that run counter to the investigator's theory of the case (David, Rawls, & Trainum, 2018). Through a disbelief repeat, investigators can make a move in this game that does not violate the rules in the sense of being overly aggressive and coercive. Focusing on how investigators respond to perceived deception allows for an analysis how such reactions can shape the manner in which the interaction proceeds.

Conclusion

Currently, there is increasing attention being paid to law enforcement interrogation techniques (see Trainum, 2016), with the growth and use of technology having contributed to this attention. Interrogations are increasingly being recorded, allowing a greater inspection of tactics and strategies used to elicit confessions (Bang, Stanton, Hemmens, & Stohr, 2018). Further, public portrayal of cases, such as the documentary television series *Making a Murderer*, have provided a window into aggressive police interrogation techniques which are reported to lead to false confessions (Gudjonsson, 1992; Gudjonsson and Pearse, 2011; Leo, 2009; Leo & Drizin, 2010; Meissner & Russano, 2003; Perillo & Kassin, 2010; Trainum, 2016).

There exists increasing professional concerns and attention over the psychologically manipulative elements and aggressive tactics of the Reid technique. For example, Wicklander-Zulawski & Associates, a major consulting and training organization—with a vast, public list of private sector, federal, law enforcement, and international clients—offers interrogation and interview training; however, in 2017 they announced that they will no longer offer training in the Reid technique, opting instead for less confrontational approaches. President and CEO Shane Sturman in a Wicklander-Zulawski press release said, "Confrontation is not an effective way of getting truthful information." Additionally, the company asserts, "The high risk of false confessions, potential for incorrect or unreliable information, and ultimately the misapplication of confrontational techniques are all reasons why WZ has chosen to no longer offer the confrontational approach in its course selections" (Wicklander-Zulawski, 2017).

At the same time, as we have noted, *less aggressive* does not mean the absence of coercion. The primary concern of any confession is whether or not it was given freely and is valid. Even when not being directly confrontational, subtle conversational features can impact how an interrogation unfolds and can shape the suspect's responses in a way that can be coercive. We do not claim that the instances examined here are examples of coercion. Our point is that tagging a statement as being deceptive puts the suspect in the position of either changing the story or further battling with investigators over whose versions of events is correct.

Given the power differential in the setting, this can be an uneven playing field (David et al., 2018; Shuy, 2014). Typically, *linguistic coercion* (Berk-Seligson, 2002) is thought of as intentional intimidation by police. This moves us from coercion as physical pressure to linguistic pressure, using conversational devices as a strategy to shape an interrogation toward a confession. This is in many respects the point of an interrogation: to use techniques to gain information and ultimately a confession from those who are guilty. At the same time, these goals must be accomplished within a legal framework and with respect to fairness and justice.

The goal of this chapter was to identify more subtle attempts of tagging deception, namely through the conversational strategy of disbelief repeats. Of course, the investigators are in a tough situation. This analysis raises a major question regarding interrogations: What are investigators supposed to do when they believe they are being lied to? We have seen the recommendation to move away from more aggressive interview tactics, as expressed in the PEACE Method and by Wicklander-Zulawski. Future studies should explore how perceived deception is expressed across a variety of interrogation approaches and see how the interrogation unfolds. Whether or not someone is in fact being deceptive, we can explore the conversational features of perceived deception, and how those expressions shape the outcomes of such encounters. By doing so, we hope to contribute to a broader examination of

deception beyond lie detection, and into sets of practices that position the perception and treatment of deception as an observable feature of interrogation. Finally, we seek to contribute to a broader understanding of how interrogation works, and works for whom.

References

Antaki, C. (1994). *Explaining and arguing: The social organization of accounts.* London: Sage.

Antaki, C. (2011). *Applied conversation analysis: Intervention and change in institutional talk.* New York: Springer.

Bang, B. L., Stanton, D., Hemmens, C., & Stohr, M. K. (2018). Police recording of custodial interrogations: A state-by-state legal inquiry. *International Journal of Police Science & Management, 20*(1), 3–18. https://doi.org/10.1177/1461355717750172.

Beinhart, L. (2017, March 1). All politicians lie but it's better if not mentioned. *Al Jazeera.* https://www.aljazeera.com/indepth/opinion/2017/03/politicians-lie-mentioned-170301115205078.html.

Berk-Seligson, S. (2002). The Miranda warnings and linguistic coercion: The role of footing in the interrogation of a limited-English-speaking murder suspect. In J. Cotterill (Ed.), *Language in the legal process* (pp. 127–146). New York: Palgrave Macmillan.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Brisbane, A. S. (2012, January 12). Should the times be a truth vigilante? *The New York Times.* Retrieved from https://publiceditor.blogs.nytimes.com/2012/01/12/should-the-times-be-a-truth-vigilante/.

Carter, E. (2011). *Analysing police interview: Laughter, confessions and the tape.* New York: Bloomsbury Publishing.

Carter, E. (2014). When is a lie not a lie? When it's divergent: Examining lies and deceptive responses in a police interview. *International Journal of Language and the Law, 1*(1), 122–140.

David, G. C., Rawls, A. W., & Trainum, J. (2018). Playing the interrogation game: Rapport, coercion, and confessions in police interrogations. *Symbolic Interaction, 41*(1), 3–24. https://doi.org/10.1002/symb.317.

Dando, C. J., & Bull, R. (2011). Maximising opportunities to detect verbal deception: training police officers to interview tactically. *Journal of Investigative Psychology and Offender Profiling, 8*(2), 189–202. https://doi.org/10.1002/jip.145.

DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74.

Drew, P. (Ed.). (1992a). *Talk at work: Interaction in institutional settings.* Cambridge, UK: Cambridge University Press.

Drew, P. (1992b). Contested evidence in court-room cross-examination: The case of a trial for rape. In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 418–469). Cambridge, UK: Cambridge University Press.

Drew, P., & Heritage, J. (2006). Editors' introduction. In P. Drew & J. Heritage (Eds.), *Conversation analysis* (Vol. 1, pp. xxi–xxxvii). Thousand Oaks, CA: Sage.

Edwards, D. (2006). Facts, norms and dispositions: Practical uses of the modal verb *would* in police interrogations. *Discourse Studies, 8*(4), 475–501.

Fallows, J. (2016). How to deal with the lies of Donald Trump: Guidelines for the media. *The Atlantic*. Retrieved from https://www.theatlantic.com/notes/2016/11/a-reflexive-liar-in-command-guidelines-for-the-media/508832/.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Cambridge: Polity Press.

Garfinkel, H. (2006). *Seeing sociologically: The routine grounds of social action*. In A. W. Rawls (Ed.). Boulder, CO: Paradigm Press. (Original work published in 1948).

Garrett, B. (2011). *Convicting the innocent*. Cambridge, MA: Harvard University Press.

Gerwehr, S., & Glenn, R. W. (2000). *The art of darkness: Deception and urban operations*. Santa Monica, CA: RAND Corporation. Retrieved from https://www.rand.org/pubs/monograph_reports/MR1132.html.

Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior*. Garden City, NY: Doubleday.

Gordon, N. J., & Fleisher, W. L. (2011). *Effective interviewing and interrogation techniques*. Boston, MA: Academic Press.

Granhag, P. A., Vrij, A., & Verschuere, B. (2015). Introduction. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches* (pp. xv–xviii). West Sussex, UK: Wiley Blackwell.

Greenberg, D. (2017, January 28). The perils of calling trump a liar. *Politico*. Retrieved from http://www.politico.com/magazine/story/2017/01/the-perils-of-calling-trump-a-liar-214704.

Gudjonsson, G. (1992). *The psychology of interrogations and confessions: A handbook*. Malden, MA: Wiley.

Gudjonsson, G., & Pearse, J. (2011). Suspect interviews and false confessions. *Current Directions in Psychological Science, 20*(1), 33–37. https://doi.org/10.1177/0963721410396824.

Heritage, J. (2012). Epistemics in action: Action formation and territories of knowledge. *Research on Language and Social Interaction, 45,* 1–29. https://doi.org/10.1080/08351813.2012.646684.

Heritage, J. (2013). Turn-initial position and some of its occupants. *Journal of Pragmatics, 57,* 331–337. https://doi.org/10.1016/j.pragma.2013.08.025.

Hess, J. E. (2010). *Interviewing and interrogation for law enforcement*. New Providence, NJ: Matthew Bender & Company Inc.

Holan, A. D. (2015, December 11). All politicians lie. Some lie more than others. *New York Times*. Retrieved from https://www.nytimes.com/2015/12/13/opinion/campaign-stops/all-politicians-lie-some-lie-more-than-others.html.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2005). *Criminal interrogation and confessions: Essentials of the Reid technique*. Boston, MA: Jones and Bartlett Publishers.

Jefferson, G. (1972). Side sequences. In D. N. Sudnow (Ed.), *Studies in social interaction* (pp. 294–233). New York, NY: Free Press.

Kessler, G., Rizzo, S., & Kelly, M. (2018, May 1). President Trump has made 3,001 false or misleading claims so far. *The Washington Post*. Retrieved from

https://www.washingtonpost.com/news/fact-checker/wp/2018/05/01/president-trump-has-made-3001-false-or-misleading-claims-so-far/.

Komter, M. (2003). The interactional dynamics of eliciting a confession in a Dutch police interrogation. *Research on Language and Social Interaction, 36*(4), 433–470. https://doi.org/10.1207/S15327973RLSI3604_5.

Leo, R. A. (2009). *Police interrogation and American justice.* Cambridge, MA: Harvard University Press. Kindle Edition.

Leo, R. A., & Drizin, S. A. (2010). The three errors: Pathways to false confession and wrongful conviction. In G. D. Lassiter & C. Meissner (Eds.), *Police interrogations and false confessions: Current research, practice, and policy recommendations* (pp. 9–30). Washington, DC: American Psychological Association.

Maynard, D. (2003). *Bad news, good news: Conversational order in everyday talk and clinical settings.* Chicago, IL: University of Chicago Press.

Meissner, C. A., & Russano, M. B. (2003). The psychology of interrogations and false confessions: Research and recommendations. *Canadian Journal of Police & Security Services, 1*(1), 53–64.

Moston, S., & Stephenson, G. M. (2009). A typology of denial strategies by suspects in criminal investigations. In R. Bull, T. Valeintine, & T. Williamson (Eds.), *Handbook of psychology of investigative interviewing: Current developments and future directions* (pp. 17–34). New York: Wiley.

Olson, E. A., & Wells, G. L. (2004). What makes a good alibi? A proposed taxonomy. *Law and Human Behavior, 28*(2), 157–176. https://doi.org/10.1023/B:LAHU.0000022320.47112.d3.

Perillo, J. T., & Kassin, S. M. (2010). Inside interrogation: The lie, the bluff, and false Confessions. *Law and Human Behavior, 35*(4), 327–337. https://doi.org/10.1007/s10979-010-9244-2.

Pomerantz, A. (1980). Telling my side: Limited access as a fishing device. *Sociological Inquiry, 50*(3–4), 186–198.

Pomerantz, A. (1984). Pursuing a response. In P. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 152–163). Cambridge: Cambridge University Press.

Public Trust in Government: 1958–2017. (2017). Retrieved from http://www.people-press.org/2017/12/14/public-trust-in-government-1958-2017/.

Rabon, D., & Chapman, T. (2009). *Interviewing and interrogation.* Durham, NC: Carolina Academic Press.

Reid Institute. (n.d.) *How the Courts view the Reid technique.* Retrieved from https://www.reid.com/newmedia/How%20Courts%20View%20the%20Reid%20Technique%20Dec%202016.pdf.

Reynolds, E. (2009). *Lies in context* (Unpublished masters thesis). Australian National University, Canberra, Australia.

Reynolds, E., & Rendle-Short, J. (2011). Cues to deception in context: Response latency/gaps in denials and blame shifting. *British Journal of Social Psychology, 50*(3), 431–449. https://doi.org/10.1348/014466610X520104.

Robinson, J. D. (2012). Epistemics, action formation, and other-initiation of repair: The case of partial questioning repeats. In M. Hayashi, G. Raymond, & J. Sidnell (Eds.), *Conversational repair and human understanding* (pp. 411–465). Cambridge, UK: Cambridge University Press.

Robinson, J. D., & Kevoe-Feldman, H. (2010). Using full repeats to initiate repair on others' questions. *Research on Language and Social Interaction, 43*(3), 232–259. https://doi.org/10.1080/08351813.2010.497990.

Sacks, H. (1972). Notes on police assessment of moral character. In D. N. Sudnow (Ed.), *Studies in social interaction* (pp. 280–293). New York: Free Press.

Sacks, H. (1975). Everyone has to lie. In B. Blount & M. Sanches (Eds.), *Sociocultural dimensions of language use* (pp. 57–80). New York, NY: Academic Press.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696–735.

Schegloff, E. A. (1991). Conversation analysis and socially shared cognition. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 150–171). Washington, DC: American Psychological Association.

Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes, 23*(3), 499–545. https://doi.org/10.1080/01638539709545001.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*(2), 361–382.

Shuy, R. W. (2014). *The language of murder cases: Intentionality, predisposition, and voluntariness.* Oxford: Oxford University Press.

Stokoe, E. (2010). 'I'm not gonna hit a lad': Conversation analysis, membership categorization and men's denials of violence toward women. *Discourse & Society, 21*(1), 59–82. http://das.sagepub.com/content/21/1/59.

Svennevig, J. (2008). Trying the easiest solution first in other-initiated of repair. *Journal of Pragmatics, 40*(2), 333–348. https://doi.org/10.1016/j.pragma.2007.11.007.

Trainum, J. (2016). *How the police generate false confessions: An inside look at the interrogation room.* Baltimore, MD: Rowman & Littlefield.

Vrij, A. (2000). *Detecting lies and deceit: Psychology of lying and its implications for professional practice.* Chichester, UK: Wiley.

Vrij, A. (2004). Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology, 9*(2), 159–181. https://doi.org/10.1348/1355325041719356.

Vrij, A., Meissner, C. A., & Kassin, S. M. (2015). Problems in expert deception detection and the risk of false confessions: No proof to the contrary in Levine et al. (2014). *Psychology, Crime and Law, 21*(9), 901–909. https://doi.org/10.1080/1068316X.2015.1054389.

Waldman, K. (2017, January 23). Lies, damned lies, and amazing new euphemisms for lying. *Slate.* Retrieved from http://www.slate.com/articles/news_and_politics/politics/2017/01/_alternative_facts_is_the_perfect_euphemism_for_the_age_of_trump.html.

Whaley, B. (1969). *Stratagem: Deception and surprise in war.* Cambridge, MA: MIT Center for International Studies.

Wicklander-Zulawski. (2017, March 6). Wicklander-Zulawski discontinues Reid method instruction after more than 30 years [Press release]. Retrieved from https://www.w-z.com/portfolio_category/wz-news/.

The History, Present, and Future of Police Deception During Interrogation

William Douglas Woody

In this chapter, I examine the history of deception in police interrogation in the United States. The chapter opens with a review of the emergence of civilian policing and the coercive interrogation tactics commonly used by police prior to 1930. Next, I explore the factors that prompted mid-twentieth-century changes, including legal investigations, journalistic and public outcry, and a series of decisions by the US Supreme Court and other courts, which in turn led to the reform of police interrogation tactics by law enforcement officers themselves. Then, I examine the current disputes and ongoing changes related to uses of police deception, including questions raised by courts, the growing awareness of false confessions, and changes promoted by those who train interrogators. The chapter concludes with an examination of legal and scientific factors that will continue to shape the theory and practice of police interrogation into the future.

Interrogation Tactics in Early Civilian Policing

Civilian policing, the use of non-military personnel to provide law enforcement services in cities and other communities, emerged relatively recently in the United States (Kelling & Moore, 1988). Between the advent of civilian policing in the United States in the mid-1800s and World War II, police interrogators relied extensively on physical coercion to gain confessions. These coercive approaches continued largely unchallenged into the 1930s, when a

W. D. Woody (✉)
University of Northern Colorado, Greeley, CO, USA
e-mail: william.woody@unco.edu

series of factors raised important questions about torture as a police interrogation tool. These factors included government investigations into police practices, journalistic and public responses, and court decisions, all of which greatly reduced the use of physical coercion and ushered in the widespread use of deceptive approaches to police interrogation in place of coercion (Leo, 1992).

Early Civilian Policing and Interrogation in the United States

Civilian policing has its first roots in Rome, when Emperor Augustus employed the elite Praetorian Guard to protect the emperor and his assets (Germann, Day, & Gallati, 1988). Much later, other models of civilian policing emerged in Western Europe that provided guidance for the United States. For example, both before and after the French Revolution (1789–1799), there were police forces in Paris and mounted patrols on French roads; however, these served less to prevent crime or protect citizens than to protect royalty and others with very high status (Elmsley, 1984). The French system, in part due to the fears it inspired in the common people, served as a foil for the British system, which in turn provided a model for the systems of policing that emerged in the United States, starting in cities on the East Coast in the 1830s and 1840s.

In the late 1700s and early 1800s, the British public observed the French system and rejected any centrally organized police force as incompatible with individual liberties (Elmsley, 1984). Additionally, there were concerns that a civilian police force would view the public in the same ways that an occupying army viewed the defeated civilian population (Elmsley, 1984). Therefore, Robert Peel, British Home Secretary in 1829 when the Metropolitan Police Act instituted civilian policing in London, established standards to assure the public that police would act for the public good rather than as occupiers. First, he eschewed lethal weapons for most officers, and he chose blue as the uniform color (which persists to this day in Great Britain as well as the United States) to distinguish police from the British Army, who dressed in well-known red coats. Second, he hired officers who lived in the jurisdictions they would police, and he set pay rates similar to those of skilled workers so that officers would be of similar financial status as most of the citizens in their districts (Critchley, 1972). Fundamentally, Peel emphasized public service (Lynam, 1964). Across decades, his endeavors slowly gained public support and served as the primary model for early policing in cities in the United States.

The United States also had its own home-grown form of civilian policing: the slave patrol (Reichel, 1988). Slave patrols emerged in the early 1700s in the Carolinas and then spread through the colonies, and they continued until the Civil War (Reichel, 1988). Slave patrols had widespread public support for their limited focus and for their explicitly racially biased enforcement goals. Despite the narrow and biased goals of slave patrols, these civilian patrols were

"a forerunner of modern American law enforcement" (Turner, Giacopassi, & Vandiver, 2006, p. 186). After the Civil War, slave patrols gave way to civilian policing in cities, but civilian police enforced Jim Crowe laws in the South and residential segregation by law in the North well into the mid-twentieth century, both in Whites-only communities—called Sundown Towns—and in cities that remained legally segregated by race (Loewen, 2005; Wilkerson, 2010; Williams & Murphy, 1990).

Civilian policing in cities in the United States emerged on the East Coast in the mid-1800s and then spread throughout the nation. Kelling and Moore (1988) refer to the era from the emergence of policing to the 1930s as the "political era" (p. 2). Typically, in larger cities, local district leaders nominated officers, and the mayor approved the officers or, in some times and places, awarded police jobs as patronage appointments. Officers served at the discretion of elected leaders with limited if any oversight from courts, journalists, or the public. There were neither educational requirements for the position nor rigorous hiring procedures. With political support for the current mayor as the most prominent criterion for hiring, departments were often corrupt (Bettman, 1974; Kelling & Moore, 1988). To keep their jobs, many officers were expected to act as enforcers for the mayor's political views, sometimes targeting people who appeared to be members of racial or immigrant minority groups with a wide range of techniques, including physical violence and intimidation (Fogelson, 1977; Leo, 1992; Oliver, 2006). These characteristics of policing directly influenced interrogation techniques.

Coercive Interrogation Tactics in the Political Era

The political era of policing brought extensive oversight by political authorities but only limited oversight by other constituent groups, such the courts, journalists, and the wider public, and common interrogation tactics reflected these facets of civilian policing. Physical coercion remained widespread, unregulated, and encouraged, "sometimes employing methods that left external signs of abuse, but more commonly using physical force in ways that did not" (Leo, 1992, p. 38). Typically called "the third degree" (Kidd, 1940, p. 45; Leo, 1992, p. 38; see Bunn, 2007; Chafee, Pollak, & Sterns, 1969), these included physical coercion (e.g., deprivation of food and water, beatings, the sweatbox, and waterboarding, sometimes called "the water cure" in this context; Leo, 1992, pp. 38–40) and psychological coercion (e.g., holding the suspect indefinitely without charge, explicit threats of harm for failure to confess or explicit promises of leniency in exchange for confession). Today, these tactics raise important concerns even when applied to those suspected of terrorist attacks (Brennan et al., 2014; Hoffman et al., 2015; Senate Select Committee on Intelligence, 2013). In the political era, however, "these tactics appeared to many as a normal, if not inevitable aspect of policing" (Leo, 1992, p. 48).

Grogan and Woody (2016) review an illustrative example from Chicago in 1906 (see also Christison, 1907). A young man who appeared to have a cognitive disability confessed to murder, and few doubted his confession. Hugo Münsterberg (1908), however, questioned whether a flash of light the suspect saw immediately before his confession led to "autohypnotisation" (p. 160), which in turn caused a false confession. According to Woody (2016), however, no observers at this time raised questions about the source of the flash: The flash was a reflection from the loaded handgun the officer pointed at the suspect in an explicit threat to kill the suspect if he did not confess. It appears nearly inconceivable today that a defense attorney (or district attorney, trial judge, appellate court, or legal journalist) would watch interrogation video depicting an officer explicitly threatening a suspect with death and then fail to raise questions about risks of false confession. But, given the coercion typically expected during police interrogations in 1906, no one did (Woody, 2016).[1] Coercive tactics persisted into the 1930s, and then change resulted from a series of factors, including legal investigations, journalistic investigations, and court decisions, all of which inspired police and others to reform policing from the inside. These reforms in turn led away from coercion and into widespread uses of deception in police interrogation.

Reforms to Police Interrogation

Kelling and Moore (1988) call the 1930s to the 1970s the "Reform Era" of policing, and reforms extended to police interrogation tactics. Several factors drove the impetus for reforms. First, legal investigations ordered by President Hoover in the 1920s led to the Wickersham Commission Report, including the Report on Lawlessness in Law Enforcement (1931), which detailed coercive interrogation tactics and reported their widespread use in police departments in the United States (Keedy, 1937; Leo, 1992). In its review of abusive interrogation practices, the report also noted the potential for these tactics to induce false confessions. Second, journalists disseminated the findings of the Wickersham Commission, producing books (e.g., Hopkins, 1931; Lavine, 1930) and several articles in national newspapers (e.g., *The Nation*, 1922, 1930; *The New Republic*, 1924, 1930) and magazines (e.g., Chafee, 1931) that publicized police abuses and stirred public ire (Bunn, 2007; Leo, 1992). These reports led to investigations of police practices during these decades (Walker, 1980) and provided support for judicial review.

A third drive for reform came from a series of early to mid-twentieth-century court decisions related to police interrogation practices. Progress on this front was neither continuous nor smooth. For example, despite the US Supreme Court's 1897 rejection of confessions generated from explicit threats and promises (*Bram v. United States*, 1897), these tactics persisted

in police departments through the early twentieth century, as noted previously in the 1906 Chicago homicide case. Despite these and other difficulties, courts began raising questions about police interrogation techniques. For example, in 1919 Ziang Sung Wan confessed to homicide. The court admitted his confession to trial, and an appellate court upheld his confession. The US Supreme Court overturned the conviction due to concerns about coercion: Police had held Wan incommunicado during nine days of interrogation in a hotel room, despite Wan's substantial physical illness, exhaustion, and emaciation (*Ziang Sung Wan v. United States*, 1924). This and other cases set the stage for *Brown et al. v. Mississippi* (1936).

Brown et al. v. Mississippi (1936) provides an illuminating and tragic glimpse into courts' typical views of interrogation tactics at this time as well as into the experiences of defendants who were not White. Four African-American suspects with limited education confessed to murder and were convicted at trial, despite discussion in court of the physically abusive police interrogation tactics. Police had induced the confessions with simulated executions (i.e., hanging until a suspect approached death), whippings, severe beatings of restrained suspects with leather belts that had metal buckles, and explicit threats of death (*Brown et al. v. Mississippi*, 1936, pp. 282–284). The defendants later attempted to retract their confessions, but despite their mistreatment, their defense attorneys did not even seek to suppress the confessions, the trial judge admitted their confessions to court, and the Mississippi Supreme Court affirmed the conviction. By overturning the Mississippi Supreme Court's decision, this became the first case in which the US Supreme Court ruled that confessions generated by torture were no longer admissible in court (Wakefield & Underwager, 1998). The US Supreme Court later took additional steps to curb specific forms of physical mistreatment of suspects, for example, the withholding of sleep or food (e.g., *Ashcraft v. Tennessee*, 1944; *Reck v. Pate*, 1961, respectively).

Police reform of interrogation tactics formed a small piece of larger reform. Though inspired by factors outside of law enforcement, several important figures in law enforcement led to these changes. August Vollmer and J. Edgar Hoover, among many others, worked to emphasize police professionalism and to reduce physical coercion during interrogation (Bunn, 2007; Kelling & Moore, 1988; Leo, 1992; Parker, 1972). These and other reformers separated police from local politics, reduced patronage, and emphasized merit in selection and promotion of officers, who now expected to retain their positions even with the election of a new mayor (Kelling & Moore, 1988). Additionally, these reformers and others worked to improve the educational requirements, moral standing, and reputations of police officers and departments (Kelling & Moore, 1988), leading to what Leo (1992) called the "triumph of professionalism" (p. 47). How did these changes affect police interrogation techniques?

Coercion to Deception

Reform of methods of police interrogation required a complex web of forces. These included legal investigations, journalism, court decisions, and the actions of important individual reformers as well as changes in public perceptions and expectations. These reforms remained slow. Although Chafee, Pollak, and Sterns (1969) report that physical coercion persisted in police departments well after World War II, Leo (2004) reports that it had largely disappeared by the mid-1960s.[2] What replaced physical coercion?

Several reformers led the move away from physical coercion, including the authors of a new wave of interrogation manuals that emphasized trickery and deception over physical torture. The first manual, by W. R. Kidd (1940; Kamisar, 2008), explicitly rejected the use of coercion. There exist similar statements in manuals by Clarence D. Lee (1953, a protégé of August Vollmer), Arther and Caputo (1959), and others. The most powerful influence, however, came from the writings of Fred E. Inbau (1942, 2004), who eschewed coercion and, as noted in his obituary, "helped elevate trickery and deceit to a high art of police interrogation" (Thomas, 1998, para. 1).

In response to the 1929 Valentine's Day murders in Chicago, Inbau, a law professor at Northwestern University, assigned John E. Reid, at the time a young lawyer, to interrogate suspects using the polygraph to detect deception (Slowik, 2016). Inbau reportedly observed many hours of interrogations and polygraph evaluations of suspects, and he tracked suspects' verbal and nonverbal behaviors. He and Reid then formalized their tactics into what is now called the Reid technique (Slowik, 2016). In this way, Inbau, Reid, and others led the move from physical coercion, with its concomitant legal and journalistic investigations, emerging limits from courts, and rising public outcry, to deception, which at the time constituted the moral high ground (Leo, 1992).

THE MID- TO LATE TWENTIETH CENTURY: DECEPTION AS THE FOUNDATION

Inbau and others argued strongly against coercive interrogation tactics (Inbau, 1942; Inbau & Reid, 1967) and emphasized deception in place of coercion, a transition that would be viewed more positively by legal investigators, journalists, courts, and the public. Additionally, Inbau (1942) emphasized deception as a more reliable approach, and he encouraged officers to employ a wide variety of tactics, including implicit threats and promises, acting as a friend of the suspect, and, most controversially, using false-evidence ploys (FEPs), false claims to have incriminating evidence (e.g., Leo, 2008; Woody & Forrest, 2009; Woody, Forrest, & Stewart, 2011).

Inbau consistently emphasized remaining within legal limits established by courts. For example, even in his early writings (e.g., Inbau & Reid, 1967), he strongly cautioned police to follow *Bram v. United States* (1897) carefully

and to avoid explicit threats or promises. He also recommended, however, careful use of *implicit* threats and promises to suspects to induce confessions while staying within the letter of the law, and he and Reid provided extensive detail and many examples to enable police interrogators to navigate this vague legal ground (see Inbau & Reid, 1967, pp. 187–195). More importantly, Inbau and Reid emphasized trickery and deception in multiple ways, but cautioned that "the trickery or deceit must not be of such a nature as to induce a false confession" (1967, p. 195). Despite this warning, Inbau and Reid provide many possible methods of deception, along with the claim, rooted in the mid-twentieth century, that "no case has prohibited [the] usage [of deception]" (1967, p. 196). Additionally, Inbau and Reid argued that deception leads to true but not false confessions, and they place deception, particularly FEPs, at the heart of their methods, claiming that "without some elements of 'trickery,' such as leading the suspect to believe that the police have some tangible or specific evidence of guilt [i.e., an FEP], many interrogations will be totally ineffective" (pp. 196–197).

Through this time, courts largely shared Inbau's views about deception. In these contexts, court approval of interrogation tactics means, as stated by Kassin (2010) "that the confessions [that these deceptive techniques] produce are admitted into evidence [at trial]" (p. 233) and upheld on appeal. Across several cases, courts examined deception during interrogation and found deception not to be inherently coercive. Among other deceptive tactics, courts have accepted a confession to an officer pretending to be a fellow inmate (*Illinois v. Perkins*, 1990), a confession to murder that resulted from questions about a burglary when officers did not tell the suspect he would also be questioned about homicide (*Colorado v. Spring*, 1987), and a confession to violence that followed deception about the victim (i.e., police falsely claimed the victim was still alive, *State v. Cooper*, 1974).

As noted previously, deception about evidence was foundational for Inbau and others, and has been largely accepted by courts (Inbau & Reid, 1967). Inbau and his colleagues argued strongly that FEPs lead to true but not false confessions. As they stated in their most recent manual, "Would this false statement [about non-existent evidence] cause an innocent person to … confess? Of course not!"; they then refer to the possibility that an FEP alone would cause a false confession as "absurd" (Inbau, Reid, Buckley, & Jayne, 2011, pp. 351–352). Courts have generally agreed, particularly in a series of decisions in the mid-1900s. The US Supreme Court has upheld a confession generated by a FEP (*Frazier v. Cupp*, 1969), and other courts followed this trend (e.g., *People v. Lira*, 1981; *State v. Cobb*, 1977; *Ward v. State*, 1980).

Despite these claims and the general acceptance of courts, FEPs remain "Perhaps, the most controversial tactic permissible within [the Reid technique of interrogation]" (Perillo & Kassin, 2010, p. 327). FEPs "have been implicated in the vast majority of documented false confession cases" (Kassin et al., 2010, p. 12). Additionally, in the carefully controlled conditions of experimental studies, FEPs substantially increase false confession rates

(e.g., Kassin & Keichel, 1996; Perillo & Kassin, 2010; Stewart, Woody, & Pulos, 2018). This disconnection has led to some of the most intense disagreements in psychology and law in the late twentieth century: There exists a strong body of scholarship demonstrating that FEPs increase false confession rates even as courts continue to accept confessions generated by FEPs (for a review, see Woody et al., 2011). Among other issues, these disputes have raised questions about the roles of experts (typically scholars of interrogation science who provide expert testimony to educate courts) who may testify during trials about deceptive interrogation tactics and the consequences of these techniques and the duties of the jury to evaluate interrogation techniques and confession evidence (e.g., Citron & Johnson, 2006; Fulero, 2010; Woody & Forrest, 2009; Woody et al., 2011). Experts who discuss the science of interrogation may find their testimony about scientific findings related to FEPs in apparent conflict with court precedents about these tactics.

How confident were Inbau, Reid, and their peers that deception does not lead to false confession? In the fourth edition of their manual (Inbau et al., 2001), they review and recommend a wide range of deceptive techniques, and they also encourage police interrogators to go beyond their text and to develop new and unique forms of interrogation deception. How should police officers evaluate the potential impacts of an untested form of deception on suspects? As recommended by Inbau et al. (2001), "A guideline that an interrogator may use in any case situation where he may be in doubt as to the permissibility of any particular type of trickery or deceit is to ask himself[3] the following question: Is what I am about to do, or say, apt to make an innocent person confess? If the answer to the question is 'no,' the interrogator should go ahead and do or say what was contemplated. If the answer is 'yes,' the interrogator should refrain from doing or saying what he had in mind" (Inbau et al., 2001, pp. 486–487; see Gohara, 2006 for additional review).[4] This confidence reflects the beliefs of Inbau et al. (2001, 2011) that deception does not cause false confessions and that courts would and should accept confessions caused at least in part by deception.

Despite Inbau's (1942) emphasis on deceptive tactics, his fears of the consequences of limitations imposed by "the new warnings" (i.e., *Miranda* warnings, Inbau & Reid, 1967, p. 195; see also Inbau, 1976), and his mid-twentieth century claims that courts have not prohibited the usage of deception during interrogation, courts have imposed some limitations on deception. Since *Blackburn v. Alabama* (1960), in which the US Supreme Court stated that psychological tactics could be coercive, a small number of relevant cases have emerged that narrowly limited the uses of deception during interrogation. In some cases, courts regarded the deception as too extreme and therefore as inherently coercive. For example, in the early 1960s, a woman confessed only after police falsely informed her that if she did not confess she would lose both her government benefits and the custody of her children (*Lynum v. Illinois*, 1963). Her confession was admitted to trial and

contributed to her conviction. After multiple appeals, the US Supreme Court viewed the confession as coerced and overturned the conviction (*Lynum v. Illinois*, 1963; see also *Spano v. New York*, 1959). Courts have also applied other narrow limits to deception, particularly about evidence. In *Florida v. Cayward* (1989), the Florida District Court rejected the defendant's confession because the fabricated evidence presented to him by police during his interrogation was identical to actual evidence (i.e., the police presented false laboratory evidence using their department's actual laboratory evidence forms). The court noted that the fabricated evidence could be mistaken for real evidence by defense or prosecuting attorneys, trial or appellate courts, or the media. Therefore, the court ruled that this tactic "offends our traditional notions of due process of law" and constitutes coercion (*Florida v. Cayward*, 1989, p. 974; see *State v. Chirokovskcic*, 2004; *State v. Patton*, 2003 for similar outcomes from New Jersey Appellate Courts). These and other limitations have had only narrow effects, leading Magid (2001) to argue "The Court has repeatedly declined the opportunity to place any specific limits on the use of deception during interrogation" (p. 1176). The mid- and late-twentieth century was a time of few limits on police deception.

The Present and into the Future

Although widespread uses of police deception had strong support from courts and trainers of interrogators in the mid-twentieth century, public and scholarly perspectives about police deception have changed substantially since Inbau first promoted these views. As one police reformer told Starr (2013), "'I think the Reid Technique was a child of its time' … But science has moved on" (para. 52). What factors are emerging today that are similar to and different from the factors that led police interrogators from coercion to deception? As in the early twentieth century, we see increased focus from the media, changes from those who train interrogators, and emerging court cases. Unlike the early twentieth century, however, there is also a substantial influence on police interrogation practices from academic research.

Media

As in the early twentieth century, recent reporters have exposed miscarriages of justice that resulted from false confessions and have questioned police interrogation practices. Specific false confession cases, such as the case of Jeff Deskovic, have garnered national attention (Balko, 2014; Innocence Project, 2018a), and some journalists have delved more deeply into police interrogation practices. Among others, Douglas Starr (2013, 2015, 2016) has asked difficult questions about deception in police interrogation rooms and other interrogation situations (e.g., schools). Starr has also inspired public discussion about the potential for false confessions.

In addition to these traditional sources, some novel media options now exist that were not available in the early twentieth century, including documentaries illuminating police interrogation. Some documentaries examine cases of proven false confessions that result from deception and other tactics, such as those of the Norfolk Four—four sailors who served extensive prison sentences after they falsely confessed to a sexual assault and homicide that they did not commit (Bikel, 2010; Wells & Leo, 2008). In another example, the exoneration of the African-American and Latino teenagers who falsely confessed to the attack in the Central Park Jogger case has generated extensive news and documentary coverage (e.g., Weiser, 2014; Burns, McMahon, & Burns, 2013, respectively). Some documentary makers have raised questions about police interrogation and the risk of false confession even when the ground truth remained unknown, as in the Netflix documentary *Making a Murderer* (Demos et al., 2015). Another documentary, *Scenes of a Crime* (Babcock & Hadaegh, 2011), evaluated the interrogation and confession of Adrian Thomas, who, since the release of the film, has had a second trial—this time without his confession—that led to his acquittal, as discussed later in this chapter. This and related publicity has the potential to raise awareness and to shift public opinion about deceptive police interrogation tactics and false confessions (see Mindthoff et al., 2018).

Interrogation Trainers

As in the early twentieth century, police and those who train them have endorsed changes. John E. Reid and Associates, the business that teaches the interrogation methods pioneered by Inbau and Reid, has made several changes to their curriculum. Through time, their training materials have contained increasing emphasis on the risk of false confessions. In 2005, their *Essentials of the Reid Technique* included an Appendix about false confessions (Inbau, Reid, Buckley, & Jayne, 2005), and their 2011 fifth edition of *Criminal Interrogation and Confessions* included a chapter on methods to distinguish between true and false confessions.[5] More recently, they dedicated a monthly investigator tip (i.e., an email update to interrogation practitioners) to the risks of false confessions (John E. Reid and Associates, 2015). In addition to other changes and growing cautions, in 2014, John E. Reid and Associates recognized and explicitly warned police interrogators about the potential consequences of excessive deception: "We [police interrogators and those who train them] are in a position to … potentially lose the future ability to misrepresent information to a suspect" (John E. Reid and Associates, 2014, para. 11).

Other trainers of police interrogators have taken similar steps. In an action that surprised many scholars of interrogation and confession, Wicklander-Zulawski & Associates recently ceased to teach the Reid technique. After more than 30 years of teaching these methods with the approval of

John E. Reid and Associates, Wicklander-Zulawski & Associates (2017) moved away from confrontational techniques, including the Reid technique (see Inbau et al., 2011). The response from John E. Reid and Associates (2017) was scathing, calling training from Wicklander-Zulawski & Associates "stale and out dated [sic]" (para. 4) and accusing Wicklander-Zulawski & Associates of promoting "a false narrative" (para. 1) about John E. Reid and Associates. Although these are corporations, the intensity of the dispute feels personal to at least some observers.[6] Wicklander-Zulawski & Associates also affirmed their commitment to incorporating academic scholarship into their training, which, as discussed below, raises important questions about police deception.

In addition to these trends, there exist options for police interrogators that avoid deception completely. Milne and Bull (1999) recommend widespread adoption of the nondeceptive and nonconfrontational PEACE approach used in the UK. Meissner and colleagues (2010, 2014, 2017) have developed and evaluated nondeceptive information gathering methods, which reduce the risk of false confessions in comparison with accusatory or confrontation methods. Several scholars (e.g., Kassin et al., 2010; Leo, 2008; Woody, Forrest, & Yendra, 2013) have called for the end of police deception during interrogation for these and other reasons reviewed below.

Scholarship

Unlike the early twentieth century, there now exists a robust body of scientific knowledge about interrogation and confession, and an explosion of research has occurred in the past decade (Kassin, 2016).[7] Much of this literature justifies questions about deception during police interrogation. For example, as noted previously, across many controlled simulation studies and a meta-analysis, FEPs increase false confession rates (e.g., Kassin & Keichel, 1996; Perillo & Kassin, 2010; Stewart, Woody, & Pulos, 2018). Additionally, similar to police detectives, jurors recognize deception and view it as coercive, but largely fail to reject confessions generated by deception (Woody & Forrest, 2009; Woody et al., 2013). To complicate matters further, judges are human decision-makers whose decisions incorporate confession evidence, even when judges perceive the interrogation as coercive and report that they rejected the confession (Wallace & Kassin, 2012; cf. Kassin & Sukel, 1997). These and other findings led credence to emerging concerns from courts.

Courts

Court precedents set the boundaries for police interrogation tactics, and therefore, courts may have the largest influence on what is considered the acceptable practice of interrogation. Unlike the mid- to late-twentieth century, courts have recently raised important questions about police deception,

and in some cases, courts have removed the qualified immunity that protects individual police officers from lawsuits (Bandler, 2014a, 2014b; *Sanchez v. Hartley*, 2016). These emerging decisions have broad consequences.

Courts have recently reconsidered previously accepted deceptive tactics. For example, police falsely informed Adrian Thomas that his infant son would die unless Thomas confessed in detail about how he had intentionally injured his son; tragically, his son was already dead (*People v. Thomas*, 2014). The appellate court ordered a new trial in which the jury did not hear Thomas's confession; he was then acquitted. Similarly, police informed Paul Aveni that an acquaintance would die unless Aveni told police what she had ingested; she had already died (*People v. Aveni*, 2012; McKinley, 2014). When Aveni immediately described the drugs he had injected into her, police then used Aveni's statements to prosecute him for criminally negligent homicide (McKinley, 2014). As in the Thomas case, the New York appellate court ordered a new trial. In these cases, deception that had been accepted by trial courts and acceptable to previous courts was rejected on appeal. It remains to be seen whether these trends will continue and lead to a substantial change on the acceptability of deception in much the same way that there was a substantial change in the acceptability of coercion. These changes are particularly relevant given the changing liability risks for individual officers.

Another ongoing series of changes has involved officers' risks of personal liability. Officers who perform their duties in good faith are generally protected from lawsuits by qualified immunity if their behavior does not violate law or an individual's constitutional rights (Schott, 2012). Rarely have officers faced consequences for deception during interrogation, but some recent court cases have increased these risks. For example, after a polygraph examination that took several hours, 16-year-old Jeff Deskovic falsely confessed to homicide while sobbing on the floor of the interrogation room (Innocence Project, 2018a). After a polygraph expert testified that the polygraph examination to which the officers subjected Deskovic was a scientific FEP (i.e., a false claim that the scientific machine had identified Deskovic as deceptive and that he should therefore confess), rather than a deception detection tool, several consequences emerged for the officer (Bandler, 2014a, 2014b). First, the court denied the officer's request for summary judgment based on his qualified immunity (*Deskovic v. City of Peekskill*, 2012; Innocence Project, 2018a). Second, a federal civil jury awarded Deskovic $40 million. This substantial award, though limited by a pretrial agreement to $10 million (Bandler, 2014b), reflects the severe consequences of Deskovic's mistaken conviction: 16 years in prison despite his innocence *and* the actual perpetrator remained free and committed another homicide while Deskovic was imprisoned (Innocent Project, 2018a).

The tragedies of Deskovic's conviction, including the growing risk of officers losing their qualified immunity due to their actions during

interrogations, combine with courts' emerging expectations that police should recognize and reject false confessions, despite the cognitive and other barriers to doing so (see Woody, 2017, for a review). In a recent case in the 10th Federal District (*Sanchez v. Hartley*, 2016), police made errors in corroborating Sanchez's confession to burglary. Their errors were not minor. Among other mistakes, a survivor described the perpetrator as a 190-pound man in his 40s with brown hair and no tattoos, and police detained a 130-pound 19-year-old with red hair and prominent tattoos covering both arms (Mitchell, 2016, *Sanchez v. Hartley*, 2016). The district attorney eventually dropped the charges against Sanchez nearly three years after the false confession, but Sanchez pursued civil action for violation of his Fourth Amendment rights (*Sanchez v. Hartley*, 2016). There exist several other cases with similar corroboration errors (e.g., Damon Thibodeaux [Innocence Project, 2018b], the teens who falsely confessed to the Central Park Jogger attack [see Garrett, 2010], Juan Rivera [Martin, 2011]) that did not lead to loss of officers' qualified immunity. The *Sanchez* case, however, establishes precedent to remove officers' qualified immunity for their actions and decisions in interrogation settings.[8] The risks to individual police officers have substantially increased.

Conclusions

Currently, several factors combine to change officers' uses of deception during police interrogation in the United States. These include growing journalistic and documentary interest in false confession stories and police interrogation tactics; changing strategies from those who train interrogators, including the development of nondeceptive alternatives; the growing body of academic research, including scholarship exploring police interrogation tactics and the likelihood of false confessions; and emerging legal precedents from court decisions. Two primary legal factors—courts' changing views of deception and increased willingness to remove officers' qualified immunity—may become the most tangible guides for future interrogation practices. What consequences will these factors have on police interrogators?

Similar to the substantial transition from coercion to deception, police in the United States appear on the cusp of a monumental change from extensive reliance on deception to embracing nondeceptive techniques. As in the mid-twentieth century, law enforcement officials and trainers have the opportunity to lead and guide this transition. I join other scholars in calling for an end to police deception (e.g., Kassin et al., 2010; Leo, 2008; Woody et al., 2013). For the sake of defendants who face risks of false confessions, the communities in which perpetrators continue to roam as police pursue false confessors, the officers who risk their finances and freedom during interrogation, and for the integrity of our criminal justice system as a whole, the time for the transition has come.

Notes

1. As another important point, Woody (2016) observed that contemporary scholars who have reviewed this event focus on Münsterberg's self-hypnosis explanation rather than the direct threat that accompanied the source of the flash (e.g., Dalby, 2014; Kassin, 2016; Kassin et al., 2010; Kassin & Gudjonsson, 2004; Starr, 2015).

2. Physical coercion persisted in at least some places. Larry Barksdale (2012) reports that when he was a new officer in the 1960s, senior officers in his department instructed him in the use of physically coercive interrogation tactics.

3. Throughout this and other editions of the Inbau et al. manual (2001, 2011), they use almost exclusively masculine pronouns for officers and for suspects.

4. This recommendation originally appears in Inbau (1976) as a "rule of thumb" (p. 251).

5. See Kassin, Goldstein, and Savitsky (2003); Kassin, Meissner, and Norwick (2005); and Honts, Kassin, and Craig (2013), for concerns about the abilities of police interrogators and other observers to distinguish between true and false confessions made by adults or juveniles.

6. Anecdotally, in my senior-level university class, Seminar in the Psychology of Interrogation and Confession, a student compared the intensity of this online exchange to exchanges between parties in contentious divorce proceedings.

7. In 1906, Münsterberg stepped into these debates as a scholar, but he did not have access to a strong body of experimental research. Also, he faced extensive legal resistance for several reasons, some of which were of his own making (Spillman & Spillman, 1993). The widespread rejection of Münsterberg's ideas included satire from John Henry Wigmore, who parodied Münsterberg with a fictitious trial description in which Mr. X. Perry Ment was found liable for injuring the profession of law (see Woody, 2016 for a review).

8. See *Robles v. Autozone, Inc.* (2008) for an example of a corporate fraud and loss investigator who faced an individual punitive damage award for excessive deception during a fraud and loss interrogation.

References

Arter, R. O., & Caputo, R. (1959). *Interrogation for investigators*. New York, NY: Copp.

Ashcraft v. Tennessee, 322 U.S. 143 (1944).

Babcock, G., & Hadaegh, B. (producers). (2011). *Scenes of a crime*. Los Angeles, CA: New Box Productions.

Balko, R. (2014, October 24). Jeff Deskovic, wrongly convicted of rape and murder, awarded $41.6 million in damages. *The Washington Post*. Retrieved April 30, 2016 from http://www.washingtonpost.com/news/the-watch/wp/2014/10/24/jeff-deskovic-wrongly-convicted-of-rape-and-murder-awarded-41-6-million-in-damages/.

Bandler, J. (2014a, October 15). Deskovic's lie detector test was 'an interrogation,' expert testifies. *Iohud*. Retrieved September 30, 2015 from http://www.lohud.com/story/news/local/2014/10/15/deskovic-federal-civil-trial-wrongful-conviction-putnam-daniel-stephens/17323927/.

Bandler, J. (2014b, October 24). Deskovic awarded $40M in wrongful conviction case. *Iohud*. Retrieved September 30, 2015 from http://www.lohud.com/story/news/local/2014/10/23/jeffrey-deskovic-wrongful-conviction-forty-million-verdict/17798527/.

Barksdale, L. (2012, November 28). Larry Barksdale responds: Only the guilty confess to crimes. *The Jury Expert*. Retrieved April 30, 2016 from http://www.thejuryexpert.com/2012/11/only-the-guilty-would-confess-to-crimes-%E2%80%A8-understanding-the-mystery-of-false-confessions/.

Bettman, O. (1974). *The good old days: They were terrible!* New York: Random House.

Bikel, O. (producer) (2010). *The confessions*. Arlington, VA: Public Broadcasting Service. Retrieved April 30, 2016 from http://www.pbs.org/wgbh/pages/frontline/the-confessions/.

Blackburn v. Alabama, 361 U.S. 199; 80 S. Cty. 274; 4. L. Ed. 2d 242 (1960).

Bram v. United States, 168 U.S. 532 (1897).

Brennan, J. O., et al. (June 27, 2013; Unclassified and approved for release December 8, 2014). *CIA comments on Senate Select Committee on Intelligence report on the rendition, detention, and interrogation program*. Washington, DC: Central Intelligence Agency.

Brown et al., v. Mississippi, 297 U.S. 278 (1936).

Burns, K., McMahon, D., & Burns, S. (directors) & McMahon, D., Burns, K., & Burns, S. (producers). (2013). *The Central Park Five*. Arlington, VA: Public Broadcasting Service.

Bunn, G. C. (2007). Spectacular science: The lie detector's ambivalent powers. *History of Psychology, 10,* 156–178.

Chafee, Z., Jr. (1931, November). Remedies for the third degree. *The Atlantic Monthly, 148*, 621–630.

Chafee, Z., Jr., Pollak, W. H., & Stern, C. S. (1969). *The third degree: Report to the National Commission on Law Observance and Enforcement*. New York: Arno Press and The New York Times.

Christison, J. S. (1907). *The tragedy of Chicago*. Chicago, IL: Self-published.

Citron, K., & Johnson, M. B. (2006). Expert testimony regarding disputed confession evidence in federal courts: *United States v. Vallejo. Journal of Forensic Psychology Practice, 6,* 1–27.

Colorado v. Spring, 479 U.S. 564 (1987).

Critchley, T. A. (1972). *A history of police in England and Wales* (2nd ed.). Montclair, NJ: Patterson Smith.

Dalby, J. T. (2014). Forensic psychology in Canada a century after Münsterberg. *Canadian Psychology, 55,* 27–33.

Demos, M., & Ricciardi, L. (directors) & Del Deo, A., Demos, M., Dennis, L., Nishimura, L., Ricciardi, D., Ricciardi, L., & Slabaugh, J. (producers). (2015). *Making a murderer*. Los Gatos, CA: Netflix.

Deskovic et al. v. City of Peekskill, 894 F. Supp. 2d 443, 449 (SDNY, 2012).

Elmsley, C. (1984). *Policing and its context: 1750–1870*. New York: Schocken Books.

Fogelson, R. M. (1977). *Big-city police*. Cambridge, MA: Harvard University Press.

Florida v. Cayward, 522 So. 2d. 971 (1989).

Frazier v. Cupp, 394 U.S. 731 (1969).

Fulero, S. M. (2010). Tales from the front: Expert testimony on the psychology of interrogations and confessions revisited. In G. D. Lassiter & C. A. Meissner

(Eds.), *Police interrogations and false confessions: Current research, practice, and policy recommendations* (pp. 211–223). Washington, DC: American Psychological Association.

Garrett, B. (2010). The substance of false confessions. *Stanford Law Review, 62,* 1051–1119.

Germann, A. C., Day, F. D., & Gallati, R. R. J. (1988). *Introduction to law enforcement and criminal justice* (Rev. 31st printing). Springfield, IL: Charles C. Thomas.

Gohara, M. S. (2006). A lie for a lie: False confessions and the case for reconsidering the legality of deceptive interrogation techniques. *Fordham Urban Law Journal, 33,* 791–842.

Grogan, C., & Woody, W. D. (2016). Forensic psychology. In W. D. Woody, R. L. Miller, & W. Wozniak (Eds.), *Psychological specialties in context: Enriching the classroom experience with history* (pp. 430–445). Electronic book: Society for the Teaching of Psychology. Retrieved from the Society for the Teaching of Psychology web site: http://teachpsych.org/.

Hoffman, D. H., Carter, D. J., Viglucci Lopez, C. R., Benzmiller, H. L., Guo, A. X., Latifi, S. Y., & Craig, D. C. (2015). Report to the Special Committee of the Board of Directors of the American Psychological Association: Independent review relating to APA ethical guidelines, national security interrogations, and torture. Retrieved April 30, 2016 from http://www.apa.org/independent-review/APA-FINAL-Report-7.2.15.pdf.

Honts, C. R., Kassin, S. M., & Craig, R. A. (2013). 'I'd know a false confession if I saw one': A constructive replication with juveniles. *Psychology, Crime & Law, 20,* 695–704.

Hopkins, E. J. (1931). *Our lawless police: A study of the unlawful enforcement of the law.* New York: Viking Press.

Illinois v. Perkins, 496 U.S. 292 (1990).

Inbau, F. E. (1942). *Lie detection and criminal interrogation.* Baltimore: Williams & Wilkins Co.

Inbau, F. E. (1976). Legally permissible criminal interrogation tactics and techniques. *Journal of Police Science and Administration, 4,* 249–251.

Inbau, F. E., & Reid, J. E. (1967). *Criminal interrogation and confessions* (2nd ed.). Baltimore, MD: Williams & Wilkins.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogations and confession* (4th ed.). New York: Aspen Publishers.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2005). *Essentials of the Reid technique.* Burlington, MA: Jones and Bartlett Learning.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2011). *Criminal interrogations and confession* (5th ed.). Burlington, MA: Jones and Bartlett Learning.

Innocence Project. (2018a). Jeff Deskovic. Retrieved April 5, 2018 from http://www.innocenceproject.org/cases-false-imprisonment/jeff-deskovic.

Innocence Project. (2018b). Damon Thibodeaux. Retrieved April 5, 2018 from https://www.innocenceproject.org/cases/damon-thibodeaux/.

John E. Reid and Associates. (2014). An update on the use of deceptive practice during interrogation. Retrieved April 30, 2016 from http://www.reid.com/educational_info/r_tips.html?serial=20140501&print=[print].

John E. Reid and Associates. (2015). False confessions. Investigator Tip: January–February, 2015. Retrieved April 30, 2016 from http://www.reid.com/r_whats-new.html.

John E. Reid and Associates. (2017). Reid's response to W-Z's March 6, 2017 press release. Retrieved June 20, 2017 from http://reid.com/w-z-response.html.

Kamisar, Y. (2008, March 10). Torture during interrogations: A police manual's foresight. *National Law Journal*, San Diego Legal Studies Paper No. 08-021. Retrieved June 1, 2015, from http://ssrn.com/abstract=1115906.

Kassin, S. M. (2010). Afterward: Deconstructing confessions—The state of the literature. In G. D. Lassiter & C. A. Meissner (Eds.), *Police interrogations and false confessions* (pp. 231–237). Washington, DC: American Psychological Association.

Kassin, S. M. (2016). False confessions: From Colonial Salem, through Central Park, and into the twenty-first century. In C. Willis-Esqueda & B. H. Bornstein (Eds.), *The Witness Stand and Lawrence S. Wrightsman, Jr.* (pp. 53–74). New York, NY: Springer.

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior, 34,* 3–38.

Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior, 27,* 187–203.

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest, 5,* 33–67.

Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science, 7,* 125–128.

Kassin, S. M., Meisser, C. A., & Norwick, R. J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior, 29,* 211–227.

Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law and Human Behavior, 21,* 27–46.

Keedy, E. R. (1937). The third degree and legal interrogation of suspects. *University of Pennsylvania Law Review, 85,* 765–777.

Kelling, G. L., & Moore, M. H. (1988, November). *The evolving strategy of police* (Perspectives on Policing, No. 4). Washington, DC: National Institute of Justice, U.S. Department of Justice, and Harvard University.

Kidd, W. R. (1940). *Police interrogation.* Brooklyn, NY: R. V. Basuino.

Lavine, E. H. (1930). *The third degree: A detailed and appalling exposé of police brutality.* New York, NY: The Vanguard Press.

Lee, C. D. (1953). *The instrumental detection of deception: The lie test.* Springfield, IL: Charles C. Thomas.

Leo, R. A. (1992). From coercion to deception: The changing nature of police interrogation in America. *Crime, Law, and Social Change, 18,* 35–59.

Leo, R. A. (2004). The third degree and the origins of psychological interrogation in the United States. In Lassiter (Ed.), *Interrogations, confessions and entrapment* (pp. 37–84). New York: Spring Science+Business Media.

Leo, R. A. (2008). *Police interrogation and American justice.* Cambridge, MA: Harvard University Press.

Loewen, J. W. (2005). *Sundown towns: A hidden dimension of American racism.* New York: W. W. Norton.

Lyman, J. L. (1964). The Metropolitan Police Act of 1829. *Criminal Law and Criminology, 55,* 141–154.

Lynumn v. Illinois, 372 U.S. 528 (1963).

Magid, L. (2001). Deceptive police interrogation practices: How far is too far? *Michigan Law Review, 99,* 1168–1210.

Martin, A. (2011, November 25). The prosecution's case against DNA. *New York Times Magazine.* Retrieved June 20, 2017 from https://www.nytimes.com/2011/11/27/magazine/dna-evidence-lake-county.html.

McKinley, J. C., Jr. (2014, January 24). Court weighs police role in coercing confessions. *New York Times.* Retrieved June 20, 2017 from http://www.nytimes.com/2014/01/15/nyregion/court-weighs-police-role-in-coercing-confessions.html?emc=eta1&_r=2.

Meissner, C. A., Evans, J. R., Brandon, S. E., Brussano, M. B., & Kleinman, S. M. (2010). Criminal versus HUMINT interrogations: The importance of psychological science to improving interrogative practice. *Journal of Psychiatry and Law, 38,* 215–249.

Meissner, C. A., Redlich, A. D., Michael, S. W., Evans, J. R., Camilletti, C. R., Bhatt, S., & Brandon, S. (2014). Accusatorial and information-gathering interrogation methods and their effects on true and false confessions: A meta-analytic review. *Journal of Experimental Criminology, 10,* 459–486.

Meissner, C. A., Surmon-Böhr, F., Oleszkiewicz, S., & Alison, L. J. (2017). Developing an evidence-based perspective on interrogation: A review of the U.S. government's high-value detainee interrogation group research program. *Psychology, Public Policy, and Law, 23,* 438–457.

Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice.* Chichester: Wiley.

Mindthoff, A., Evans, J. R., Perez, G., Woestehoff, S. A., Olaguez, A. P., Klemfuss, J. Z., … Woody, W. D. (2018). A survey of potential jurors' perceptions of interrogations and confessions. *Psychology, Public Policy, and Law, 24*(4), 430–448.

Mitchell, K. (2016, January 12). Ruling clears the way for false arrest lawsuit against Douglas County. *The Denver Post.* Retrieved June 20, 2017 from http://www.denverpost.com/news/ci_29374134/ruling-clears-way-for-false-arrest-lawsuit-against-douglas-county.

Munsterberg, H. (1908). *On the witness stand.* New York, NY: Doubleday. Retrieved April 30, 2016 from http://psychclassics.yorku.ca/Munster/Witness/confessions.htm.

National Commission on Law Observance and Enforcement. (1931). *Report on lawlessness in law enforcement* (Vol. 11). Washington, DC: United States Government Printing Office.

Oliver, W. M. (2006). The era of Homeland security: September 11, 2001 to …. *California Law Enforcement, 40*(3), 19–29.

Parker, A. E. (1972). *The Berkeley police story.* Springfield, IL: Charles C. Thomas.

People v. Aveni, 100 A.D.3d 228, 232 (N.Y. App. Div., 2012).

People v. Lira, 119 Cal. App. 3d 837; 174 Cal. Rptr. 207 (1981).

People v. Thomas, 22 N.Y.3d 629; 8 N.E.3d 308; 985 N.Y.S.2d 193 (2014, New York).

Perillo, J. T., & Kassin, S. M. (2010). Inside interrogation: The lie, the bluff, and false confession. *Law and Human Behavior, 35,* 327–337.

Reck v. Pate, 367 U.S. 433 (1961).

Reichel, P. L. (1988). Southern slave patrols as a transitional police type. *American Journal of Police, 7,* 51–77.

Robles v. Autozone, Inc., 2008 WL 2811762 (Cal Ct. App. 2008).

Sanchez v. Hartley, 810 F.3d 750 (10th Cir. 2016).

Schott, R. G. (2012, September). Qualified immunity: How it protects law enforcement officers. *FBI Law Enforcement Bulletin*.

Senate Select Committee on Intelligence (approved December 13, 2012; Declassification revisions December 3, 2013). *Committee study of the Central Intelligence Agency's detention and interrogation program*. United States Senate.

Slowik, S. M. (2016). Interrogation. In C. R. Berger & M. E. Roloff (Eds.), *The international encyclopedia of interpersonal communication*. Hoboken, NJ: Wiley. Retrieved June 20, 2017 from http://onlinelibrary.wiley.com/doi/10.1002/9781118540190.wbeic220/pdf.

Spano v. New York, 360 U.S. 315 (1959).

Spillman, J., & Spillman, L. (1993). The rise and fall of Hugo Münsterberg. *Journal of the History of the Behavioral Sciences, 29,* 322–338.

Starr, D. (2013, December 9). The interview: Do police interrogation tactics produce false confessions? *The New Yorker*. Retrieved June 20, 2017 from http://www.newyorker.com/magazine/2013/12/09/the-interview-7.

Starr, D. (2015, March 5). Remembering a crime you didn't commit. *The New Yorker*. Retrieved June 20, 2017 from http://www.newyorker.com/tech/elements/false-memory-crime.

Starr, D. (2016, March 25). Why are educators learning how to interrogate their students? *The New Yorker*. Retrieved June 20, 2017 from http://www.newyorker.com/news/news-desk/why-are-educators-learning-how-to-interrogate-their-students.

State v. Chirokovskcic, 860 A.2d 986 (N.J.Super. 2004).

State v. Cobb, 115 Ariz. 484; 566 P.2d 285 (AZ 1977).

State v. Cooper, 217 N.W.2d 589 (IA 1974).

State v. Patton, 826 A.2d 783 (NJ 2003).

Stewart, J. M., Woody, W. D., & Pulos, S. (2018). The prevalence of false confessions in experimental laboratory simulations: A meta-analysis. *Behavioral Sciences & the Law, 36,* 12–31.

The Nation. (1922, January 25). The third degree. *The Nation, 114*, 86–87.

The Nation. (1930, October 1). Without the law. *The Nation, 131*, 339.

The New Republic. (1924, November 12). The third degree. *The New Republic*, 272–274.

The New Republic. (1930, March 19). American Torqumadas. *The New Republic,* 113–114.

Thomas, R. M., Jr. (1998, May 28). Fred Inbau, 89, Criminologist who perfected interrogation. *New York Times*, Retrieved April 30, 2016 from http://www.nytimes.com/1998/05/28/us/fred-inbau-89-criminologist-who-perfected-interrogation.html.

Turner, K. B., Giacopassi, D., & Vandiver, M. (2006). Ignoring the past: Coverage of slavery and slave patrols in criminal justice texts. *Criminal Justice Education, 17*, 181–195.

Wakefield, H., & Underwager, R. (1998). Coerced or nonvoluntary confessions. *Behavioral Sciences & the Law, 16,* 423–440.

Walker, S. (1980). *Popular justice: A history of American criminal justice*. New York: Oxford University Press.

Wallace, D. B., & Kassin, S. M. (2012). Harmless error analysis: How do judges respond to confession errors? *Law and Human Behavior, 36,* 151–157.

Ward, v. State, 408 N.E.2d 140 (Ind. App 1980).

Weiser, B. (2014, June 19). 5 exonerated in Central Park Jogger Case agree to settle suit for $40 million. *New York Times*, Retrieved June 20, 2017 from https://www.nytimes.com/2014/06/20/nyregion/5-exonerated-in-central-park-jogger-case-are-to-settle-suit-for-40-million.html.

Wells, T., & Leo, R. A. (2008). *The wrong guys: Murder, false confessions, and the Norfolk four*. New York, NY: The New Press.

Wicklander-Zulawski & Associates. (2017). Identify the truth. Retrieved June 20, 2017 from https://www.w-z.com/truth/.

Wilkerson, I. (2010). *The warmth of other suns: The epic story of America's great migration*. New York: Vintage.

Williams, H., & Murphy, P. (1990, January). *The evolving strategy of police: A minority view* (Perspectives on Policing, No. 13). Washington, DC: National Institute of Justice, U.S. Department of Justice, and Harvard University.

Woody, W. D. (2016). The rich and interactive history of psychology and the law. In W. D. Woody, R. L. Miller, & W. Wozniak (Eds.), *Psychological specialties in context: Enriching the classroom experience with history* (pp. 411–429). Electronic book: Society for the Teaching of Psychology. Retrieved from the Society for the Teaching of Psychology web site: http://teachpsych.org.

Woody, W. D. (2017). Lowering the bar and raising expectations: Recent legal decisions in light of the scientific study of interrogation and confession. *Wyoming Law Review, 17*(2), 419–455.

Woody, W. D., & Forrest, K. D. (2009). Effects of false-evidence ploys and expert testimony on jurors' verdicts, recommended sentences, and perceptions of confession evidence. *Behavioral Sciences & the Law, 27,* 333–360.

Woody, W. D., Forrest, K. D., & Stewart, J. M. (2011). False confessions: The role of police deception in interrogation and jurors' perceptions of the techniques and their outcomes. In A. E. Hasselm (Ed.), *Crime: Causes, types and victims* (pp. 1–37). New York: Nova Publishing.

Woody, W. D., Forrest, K. D., & Yendra, S. (2013). Comparing the effects of explicit and implicit false-evidence ploys on Mock Jurors' verdicts, sentencing recommendations, and perceptions of police interrogation. *Psychology, Crime and Law, 20,* 603–617.

Ziang Sung Wan v. United States, 266 U. S. 1 (1924).

Deception Induced Confession: Strategies of Police Interrogators and Their Lay Collaborators

Tyler N. Livingston, Peter O. Rerick,
J. Guillermo Villalobos and Deborah Davis

Aspiring to become the master interrogator of his generation, a new US law enforcement officer walks into the office of the reigning master interrogator of his day to ask his advice. The master interrogator tells him:

> Son, you're trying to convince your targets to do something that they firmly believe will have terrible consequences for them, and maybe their whole group or country. They have really good reasons to believe this. And indeed, if they confess or provide intelligence, it will actually be terrible for them and others they might be protecting. They aren't going to start out liking you or trusting you. They'll assume you're the enemy and that you just want to hurt them. All their suspicions are right. They *should* fear and distrust you, and they shouldn't follow any of your advice. So, how do you get past that? Well, you're going to

T. N. Livingston · P. O. Rerick · J. G. Villalobos · D. Davis (✉)
University of Nevada, Reno, NV, USA
e-mail: debdavis@unr.edu

T. N. Livingston
e-mail: livingston@nevada.unr.edu

P. O. Rerick
e-mail: prerick@nevada.unr.edu

J. G. Villalobos
e-mail: jvillalobos@unr.edu

have to lie convincingly about everything. You've got to be the best damn liar they've ever met!

And that—along with some detailed recommendations about what lies to tell and how to tell them—has represented the essence of law enforcement interrogation training in the US for decades. We begin this chapter by describing basics of interrogation practices and the deceptions they entail. We then move to deceptions entailed in two forms of police interrogation through use of lay collaborators: covert recorded phone calls by lay parties coached to elicit admissions from suspects (pretext calls) and conspiring snitches coached to elicit confessions.

Interrogation as Psychological Manipulation: How We Got Here

Because suspects are typically hesitant to make self-incriminating statements, interrogators have historically attempted to overcome their resistance through a variety of means. From the late nineteenth century until the 1930s, the law in the US permitted police to employ "third-degree" interrogation methods including various forms of physical violence, deprivation of basic needs such as food and sleep, and explicit threats of further physical harm. These methods, since declared illegal in the US, often produced coerced false confessions from suspects who capitulated in order to escape torturous circumstances (for reviews see Davis & Leo, 2014; Leo, 2008). One commentator estimated that police during this time solved approximately 70% of criminal cases via third-degree methods (Lavine, 1930, 1936).

Because third-degree tactics led to high rates of false confession, the law changed in the US during the twentieth century to afford new legal protections to criminal suspects. Lawmakers in the 1930s prohibited interrogators from engaging in acts of violence or making explicit threats and promises contingent on confession, and the US Supreme Court eventually ruled in *Miranda v. Arizona* (1966) that criminal suspects could lawfully elect not to speak with interrogators. In response to such modern protections of criminal suspects, interrogators in the 1930s began to develop sophisticated methods of psychological deception and manipulation as primary means of obtaining confessions (for reviews see Davis & Leo, 2014; Leo, 2008).

Current Interrogation Practices

In its current dominant form, police interrogation in the US relies heavily on trickery and deception to convince the suspect of a very, very big lie: that confessing is the best way to achieve the best possible legal outcomes given the circumstances. In reality, confession can only be wise when there is a deal *in writing* regarding the suspect's outcomes. In the absence of such

an explicit deal contingent on confession, the confession will only result in a series of disadvantages and poorer outcomes for the suspect: ranging from the nature of charges filed through any post-conviction appeals (see Leo & Davis, 2010 for review).

To sell this overarching lie, interrogators rely on the combination of misdirection, a series of specific instrumental lies, and well-validated psychological weapons of social influence (Davis, 2010; Davis & O'Donohue, 2004). Interrogation training in the US is conducted primarily by professional training organizations such as John E. Reid & Associates, Inc. (reid.com), Wicklander-Zulawski & Associates, Inc. (w-z.com), and secondarily by the FBI and by local police departments. Until recently, all have trained using the Reid method or close variations thereof. Interrogation techniques are also detailed in several prominent interrogation manuals and supporting materials: Many sold at the Reid and Wicklander-Zulawski websites, as well as Amazon and other booksellers. For decades, Reid and Wicklander-Zulawski have offered training seminars throughout the US weekly and claim to have trained hundreds of thousands of law enforcement officers at all levels.

Based on these materials, the FBI and local trainers have produced workbooks used for their own training seminars. The FBI has unquestionably trained using the Reid technique in the past, as indicated by their own training materials and by numerous articles posted online in the FBI Law Enforcement Bulletin (http://www.fbi.gov/stats-services/publications/law-enforcement-bulletin). This has changed, however, as the FBI is now leading efforts toward interrogation reform. Wicklander-Zulawski announced on March 6, 2017, that they will no longer train the Reid method because of the increased risk it poses of false confession. In place of the Reid technique, this particular training organization and many prominent interrogation scholars recommend the use of non-confrontational techniques and science-based strategies to gather accurate information about the crime rather than to pursue confession as the principal goal of interrogation (see Meissner et al., 2014 for review).

But there will be a significant lag in the US before many interrogators change their methods. Local jurisdictions remain highly likely to train the Reid technique. Thus, in the sections to come, we review the basics of this technique and the deceptions involved.

Inducing the Suspect to Talk to the Interrogator

Clearly, to elicit a confession the interrogator must induce the suspect to talk to him. Since resistance is likely to be raised once a person realizes he is to be interrogated as a suspect, the interrogator prefers to elicit as much information as possible before this happens. Interrogators aim to accomplish this goal in a variety of ways, including by failing to reveal to a suspect that he or she

is, indeed, suspected of committing a crime. Detectives sometimes invite suspects to attend a "fact-finding" interview in the police station to aid in their investigation. Suspects might come to the police station completely unaware that they are entering an interrogation. Importantly, police are not legally required to read *Miranda* warnings to people who have not been placed under arrest. Thus, suspects who believe they are helping the police conduct an investigation sometimes do not receive *Miranda* warnings until they have been led to provide a full confession (Davis, 2010).

If the interrogator must arrest the suspect, he will attempt to mislead him in several ways that reduce the likelihood the suspect will invoke his right to silence. He may say something like "I want to talk to you to get some things straightened out. But *before I talk to you* I have to get you to sign some stuff" or "We have to take care of this 'formality' *before we can talk*." In this way, the interrogator misleads the suspect by conveying the presumption that they *will* talk, rather than the message that the suspect does not have to talk. And, he "trivializes" the importance of the suspect's rights and the reason he is being reminded of them (see Scherr & Madon, 2013 for effects of this strategy).

If these strategies are successful and the interrogation continues, a series of deceptions are used to facilitate confession.

Promoting Trust of the Interrogator

Any agent of influence is more effective if liked (Cialdini, 2008). Thus, as quickly in the process as possible, the interrogator seeks to undermine resistance to himself and to his messages by making himself more likeable, and seemingly more trustworthy and sympathetic to the suspect. Often, this begins with chatting and questioning on background issues or other event-relevant people and information before the suspect is accused of committing the crime. The interrogator is friendly and chatty, emphasizing similarities (whether true or not) between himself and the suspect to enhance likeability and trust. As we discuss below, other strategies are designed to further the illusion that the interrogator is the benevolent and trustworthy ally of the suspect, one who wants to help the suspect achieve the best possible legal outcomes under the circumstances.

Selling Confession

Once the interrogator accuses the suspect of guilt and transitions from interview to interrogation, he engages in an additional set of deceptions that help him sell confession as the suspect's best option (see Davis, 2010; Davis & Leo, 2014; Davis & O'Donohue, 2004; Kassin et al., 2010 for documentation of interrogator use of the following tactics and reviews of their effects).

(1) **Inflate the Strength of Evidence**

A suspect who knows he or she is innocent and/or who believes he or she will be able to convince others of that innocence will see no reason to confess. Thus, the first broad goal of interrogation is to convince suspects that establishing innocence is not a possibility. Instead, the suspect is hopelessly implicated by the evidence and is certain to be viewed as guilty by all who evaluate his case. If this is successful, the suspect is likely to turn attention to how to minimize the consequences of his seeming guilt. This renders him more susceptible to information the interrogator offers concerning how confession might do this.

If the law forbade misrepresentation of the evidence to suspects, this tactic might be effective primarily with the guilty. But the law allows interrogators to lie about evidence in such forms as falsely claiming that the suspect failed a polygraph, claiming nonexistent forensic evidence, eyewitnesses, incriminating audios or videos, and much more. And such false evidence has been strongly implicated in the elicitation of false confessions from innocent suspects (see Kassin et al., 2010 for review).

(2) **The Interrogation is an "Opportunity" for the Suspect**

The interrogator's second goal is to mislead suspects regarding the true purpose of interrogation, which is, of course, to elicit a full and detailed confession that will secure the suspect's conviction. The more the suspect becomes convinced that evidence of his guilt appears overwhelming and incontrovertible, the more easily the interrogator can turn the suspect's attention from trying to argue his innocence to trying to achieve the best outcomes available. To facilitate this transition, the interrogator attempts to frame the interrogation as an opportunity for the suspect to achieve optimal legal outcomes, given that his guilt is established.

The interrogator claims that he does not need to be there given that guilt is established, but that he wants to give the suspect an opportunity to "help himself" by "explaining" how and why the crime occurred. If he doesn't "explain," others' assumptions or other witness accounts will presumably be thought true and those will be what is presented to the jury; and they are not as likely favorable to the suspect as his own "explanation." Either course, however, ties the suspect to the crime and will help secure his conviction.

(3) **The Consequences of Guilt Are Flexible**

These messages entail several distinct misrepresentations in addition to the fundamental deception regarding the purpose of the interrogation. Among the most crucial is the idea that even if one is guilty, the consequences are flexible and depend in some way on whether the suspect cooperates and

confesses, and on the "explanation" he gives. This idea can be true in some respects, in that details of what a suspect admits to might affect the exact nature of charges against him. But detectives tend to convey much more flexibility than actually exists.

For example, the detective might mention getting the suspect counseling or help several times during an interrogation to convey the impression that the suspect might be let go with counseling instead of criminal charges. This and other tactics designed to convey the illusion of flexible (or minimal) consequences are so effective that some suspects expect to be released with a warning, counseling, or no consequences after full detailed confessions to crimes such as rape of a child, murder, and others (e.g., Drizin & Leo, 2004).

(4) **The Interrogator is a Beneficent Authority Who Can Affect These Consequences**

Whether or not suspects entered the interrogation aware of their status as suspects, once they realize they are not simply helping the police conduct an investigation, they may view the interrogator as an adversary who should not be trusted. In anticipation of this attitude, interrogators present themselves as benevolent allies who want (and have the ability) to help suspects achieve the best possible legal outcomes.

Davis and colleagues described a set of tactics they call "the sympathetic detective with a time-limited offer" (Davis, Leo, & Follette, 2010; Davis & O'Donohue, 2004). As described earlier, these tactics can begin before the suspect is accused of the crime, when the interrogator maintains a friendly demeanor during initial background questioning. Interrogators attempt to impart a sense of personal similarity with the suspect to assist in establishing rapport and, ultimately, trust. They express their beneficence both implicitly, via a friendly demeanor, and explicitly, via statements designed to demonstrate intentions to provide help. They flatter the suspect (e.g., "I think you're a stand up guy, no predator. You're just a nice guy who made a mistake…") and express their desire to help ("I'd like to help you, Buddy…"), but contingent on confession ("…but I can't help you if you don't tell me the truth"; i.e., confess). But this offer is time-limited (see below): "I can't help you once you leave here. I'll have to turn you over to the DA, and she'll do what she's gonna do." Such statements suggest both that the detective *wants* to help the suspect and that he *can* do so. Neither is true, but both are often convincing.

The effects of perceived interrogator beneficence and authority on decisions to confess are supported empirically as well as anecdotally. Davis and colleagues (2010), for example, showed that the extent to which participants viewed the interrogator as liking and wanting to help the suspect (i.e., beneficence) and as having the authority to influence whether and which charges were filed against the suspect (i.e., authority) predicted participants' recommendations as to whether both innocent and guilty suspects should confess to achieve the best legal outcomes. That is, participants were more likely to

believe the detective's recommendation to confess was wise when he was perceived as actually wanting the best for the suspect and as actually able to help him. These relationships have since been replicated in recent research (e.g., Villalobos, Kemmelmeier, & Zimmerman, 2018).

(5) The Interrogator's Help is Time-Limited

As briefly acknowledged above, once the interrogator sufficiently convinces the suspect of his or her beneficence and ability to help, the interrogator leverages the scarcity principle of social influence by communicating to the suspect that his or her help is only available during the period of the interrogation (Cialdini, 2008; Davis, 2008). Interrogators repeatedly offer help with the caveats that the suspect must tell the truth and that once the interrogation ends, the suspect will be turned over to others who will no longer care to hear his story. This time-limited offer has the effect of focusing the suspect's attention on maintaining the interrogator's help rather than on the reality of the situation, including the severe consequences of confession.

(6) The Suspect's Ability to Help Himself is Time-Limited

Interrogators also suggest that suspects can only *help themselves* during this period. The interrogator might insist that a suspect's only opportunity to tell his or her account of events is during the interrogation. Of course, this claim is patently false, as suspects in the US have the legal opportunity to speak with and in front of friends, family, defense attorneys, judges, and juries after the interrogation. Even so, if the case goes to trial, many suspects will not elect to testify, and what they said during the interrogation might be the only words from the suspect presented to jurors. Therefore, if the suspect confesses or makes lesser incriminating statements during interrogation, he will not be able to directly contradict them before the jurors.

(7) Confession is Advantageous or Without Consequences

To convince suspects the confession is in their best interests, the interrogator uses "minimization" to lower the perceived costs of confession and "maximization" to increase perceived costs of denial (Kassin, 1997). These manipulated perceptions, of course, represent the exact opposite of the real consequences of confession.

Interrogators are not legally permitted to use explicit threats or promises to convey incentives for confession (though some do anyway). They must do so indirectly, by implication, hints, and so on. The primary way in which minimization is accomplished is through a process called "theme development" (Inbau, Reid, Buckley, & Jayne, 2013). The interrogator offers scenarios for how and why the crime occurred that seem minimally serious, and perhaps

without consequences. For example, he may suggest that the suspect stole to feed his family rather than out of greed and follow up by saying: "If that's what happened, I can understand that. You have to do what's necessary to protect your family. That's no big deal. We can work with that. But if you went out and stole this stuff so you could buy drugs or gamble, that's completely different. No one is going to want to talk to you."

Alternatively, the interrogator might suggest that what the suspect did was a mistake and that he needs help to make sure it doesn't happen again: "Look, Buddy, I think you're basically a good guy. You just got yourself in over your head and made a mistake. You need to come clean and explain this so we can get you some help and move on." Such messages give the illusion that the consequences of admitting will not be as severe as one might think, given the nature of the crime.

Maximization must also be accomplished through messages implying that the consequences of denial will be negative. Such messages suggest, for example, that judges and juries will look more favorably upon a suspect who admits guilt and expresses remorse. "Who would you look more favorably toward? The guy who stands up like a man, takes responsibility and says, 'Yeah, I made a mistake. I'm sorry,' or the absolutely remorseless guy who just lies to your face, totally refuses to take responsibility for what he knows damn well he did?" A statement like this implies that a "stand-up guy" who takes responsibility for the alleged criminal activity will receive more forgiveness (and a lighter sentence) than someone who just denies and lies to the judge and jury.

The Effectiveness of Psychological Approaches to Interrogation

The many deceptive messages of modern psychological approaches to interrogation have been highly effective at producing confessions. Empirical examinations of confession rates in the US demonstrate that approximately one-half to two-thirds of all interrogated suspects ultimately confess (Kassin et al., 2010; Leo, 1996; Thomas, 1996). One survey of investigators estimated that approximately 23% of *innocent* suspects made self-incriminating statements or gave partial or full confessions. Further, a number of compilations of cases of wrongful conviction have shown that from 13% to over 60% of the wrongfully convicted had falsely confessed or been convicted in part on the basis of others' false confessions (Kassin et al., 2007; see Kassin et al., 2010 for review). Thus, while the use of deception has been effective in eliciting criminal confessions, this has come with the price of inducing many innocent individuals to falsely confess.

Indirect Interrogation: Pretext Calls

Police are not only agents of deception. They also recruit others to engage in deceptive interrogations on their behalf. The first of these forms of indirect police interrogation is the "pretext call," sometimes referred to as a "covert" or

"controlled" call. Police solicit the cooperation of persons whom a guilty suspect would expect to have some knowledge of the crime and/or incriminating activities of the suspect, to encourage the suspect to talk freely about the crime. Police ask the alleged victim to call the suspect, supposedly for another purpose (the "pretext" for the call), while police record the call. Police give the pretext caller instructions concerning which admissions or statements to elicit from the suspect, as well as specific techniques to use: many similar to those they use themselves. Though most such deceptive interactions take place in calls, there are also pretext text messages and emails, and sometimes recorded personal interactions.

Pretext calls take place in many types of criminal cases. In some cases, another person involved in the criminal activities has agreed to cooperate in exchange for benefit (see discussion of "Informants" below). In this case, the suspect can be under the illusion that the conversation is just a normal discussion of the situation at hand. In other cases, the caller might be an alleged victim of the suspect (or advocate of the victim) who confronts the suspect with an accusation. For example, an alleged rape victim might call the suspect and attempt to elicit an admission that the suspect knew the sex was not voluntary, an apology, or both.

Perhaps the most important aspect of pretext calls is that recipients are unaware that calls are recorded. This can lead suspects to speak more freely and perhaps make more incriminating statements. Moreover, it replicates several desirable conditions that police aim to create in custodial interrogations. First, suspects are effectively talking to the police without an attorney present. Second, pretext calls can cause recipients to develop goals that would facilitate admissions. For example, if a tearful, distressed caller accuses the recipient of rape, the accused might feel sorry for the caller and try to manage her distress. To do so, he may offer an apology or "explanation." Third, because suspects have not yet been placed under arrest, there is no requirement to administer Miranda rights or obtain a waiver (Davis, 2010).

Pretext Calls vs. Police Interrogation: Similarities and Differences

The purposes of pretext calls and police interrogations are the same: to elicit sufficiently incriminating admissions from suspects to support relevant criminal charges and to secure a guilty plea or guilty verdict at trial. However, the extent to which common interrogation techniques are used in pretext calls depends on the type of caller.

In non-accusatory pretext calls, the caller attempts to converse in a way the suspect will perceive as entirely normal. The caller must lead the conversation in the direction of the events in question without seeming suspicious. Sometimes, the call might be framed as a business discussion of an ongoing criminal enterprise. Other times, the caller might discuss the relevant activities under the guise of voicing concerns about police investigations or thoughts about how to cover up the crime successfully. These calls typically have less in common with the techniques of police interrogation compared to accusatory calls.

Accusatory pretext calls might take place when the caller is a coconspirator turned snitch, but they are much more likely to occur when the caller is an alleged victim of a crime such as rape or child sexual assault. An associate or advocate of the victim, such as the victim's mother, can also make the call. The next sections illustrate pretext callers' use of several common police interrogation techniques in cases of rape or child sexual assault.

What Do I Want Out of This? Feigning Benevolence

Recall that in custodial interrogations, interrogators present themselves as beneficent allies to the suspect who want to help him achieve the best outcomes (Davis et al., 2010). Pretext callers can create the appearance of beneficence with relative ease. The caller can use pre-existing trust from the relationship's past, rather than having to build it during a high-stress interaction between strangers. The caller might act as if his or her interests are aligned with those of the suspect. Police coach callers to take advantage of that trust, and they do so convincingly. Notice the analogy, for example, between this pretext caller's statements and the "sympathetic detective" interrogation tactic.

> *CALLER*: "I love you. I want to be able to keep our family together, but I can't go on until you admit what you did and explain it. How can I trust you?"
> *CALLER*: "I'm not trying to get you in trouble, I would rather keep this between us. But you have to tell me the truth."

Just like the interrogator, the caller expressed sympathy and at least desire not to hurt the caller. But she also made the offer to continue the relationship contingent on the recipient admitting and explaining.

Why Are We Talking: Misrepresenting
the Purpose of the Interaction

Both police interrogators and pretext callers misrepresent the purpose of the interaction. In both cases, the suspect is firmly accused of the crime, and no denials are accepted. The interrogator's purpose is typically misrepresented as the intention to investigate why the crime was committed in order to contribute to the illusion that the consequences for the crime in question are flexible, and that the interrogator has some authority over those consequences (Davis et al., 2010). A pretext caller might also ask why the suspect committed the crime in question, but say she needs to know to help her heal or move on, or to be able to continue the relationship. This tactic serves several purposes.

First and foremost, to answer the question of why, one must admit that the action occurred. Asking *why* typically occurs in the context of the caller's unwillingness to accept a denial. This tactic helps to convince the suspect that he cannot hope to successfully claim innocence with the caller and, as in interrogation, might turn his attention to how best to "explain."

In addition, the question might be designed to manipulate the suspect's emotions. Pretext callers who are alleged victims commonly express considerable distress during the call. This expression of emotion is intended to induce guilt and motivate the suspect to try to ease the caller's distress. The latter might prompt the suspect to explain the crime, apologize, offer to sever all contact from the caller, or even admit guilt directly to the caller. All appear to imply guilt.

Among the strategies of displays of distress, the caller might tell the suspect that he or she "needs" to understand why the suspect raped or sexually assaulted him or her in order to move on, as illustrated by examples from a child sex abuse case:

> CALLER: "Why would you do this? You really hurt me, you know."
> CALLER: (crying) "Why did you do this to me? I was just a little kid. Why would you take away my childhood like this? I can't get past this when you won't even admit what you did."

Similar to the tactics utilized by police interrogators, the caller might state or imply that the suspect's explanations of why the crime was committed can affect whether the caller will end his or her relationship with the suspect or report alleged criminal activity to the police.

The caller might also explicitly ask for an apology, assuming that jurors will perceive an apology as incriminating during trial. Again, the caller might express his or her "need" for an apology, sometimes along with claims of how it will help him or her.

> CALLER: "I really, I really need you to just apologize."
> SUSPECT: "I'm so—I'm so sorry."
> CALLER: "And promise that you will never do this to anyone again."
> SUSPECT: "I promise. I promise and I—I'm so sorry."

The Importance of Admitting to Knowledge of Nonconsent

When the caller is an alleged victim of rape, a crucial admission that both interrogators and pretext callers seek is acknowledgment that the suspect *knew* at the time that the victim did not want to have sex (i.e., that the sex was nonconsensual). This request for simple acknowledgment is almost universal in pretext calls for sexual crimes, as illustrated by the following examples.

> CALLER: "You knew I didn't want to have sex with you, didn't you? I kept trying to push you away. I know you knew that! Why can't you just admit what you did?"
> CALLER: "You know, you basically kind of raped me that night, right?"

The Use of Threats and Promises to Elicit Confessions

The law in the US prohibits use of explicit threats and promises by police interrogators. Interrogators who violate these prohibitions risk the exclusion of all admissions they elicit from presentation at trial. Pretext callers face no such legal prohibitions. This freedom allows callers to utilize a variety of threats.

Most common among the legal threats are those concerning reporting to the police or to other authority figures such as parents or therapists (who might themselves report to police). The caller might also imply that he or she will report the accusation in the absence of a satisfactory response from the suspect. Or, she may just mention the idea without a threat, knowing that it would raise the suspect's fear and perhaps his compliance with her demands.

> *CALLER*: "Listen, I didn't tell them your name, I didn't tell them any of that. I'm just afraid that you're going to get in trouble and I don't know what I should do."
> *CALLER*: "It is going to become a much bigger deal. If you won't deal with this, I'll have to report it and get others involved."

Also common are relationship threats and promises. These might be directly stated as such or implied.

> *CALLER*: "I'm not coming home until you're ready to talk about it."
> *CALLER*: "If I can't trust you, I can't be with you. I don't want to sacrifice our family or our marriage."

The suspect's wife cleverly combined threats that his marriage would end if he did not confess, with promises that his confession would be kept between the two of them and the implication that counseling would be the extent of his consequences.

> *CALLER*: "If you did touch her, you need to tell me so we can get you some help."

Theme Development

Earlier we described the interrogation technique "theme development," whereby interrogators offer scenarios for how and why the crime might have occurred that are either noncriminal or do not seem very serious. Police explicitly coach pretext callers to use this strategy in some cases. For example, the caller below offers many innocuous scenarios for how her husband might have touched her daughter's vagina.

> *CALLER*: "Maybe it was an accident or something."

    CALLER: "Have you touched her private area in a way to help her clean herself?"
    CALLER: "Were you trying to teach her something about inappropriate touching, and you maybe showed or pointed to that area?"

Note that some such scenarios are difficult to discount completely. How, for example, can a person be certain that he or she never accidentally brushed a private area throughout years of handling and playing with kids in the most innocent fashion? But these scenarios also give the accused an opportunity to admit that something happened that might satisfy the caller while still seeming innocent. From the point of view of police, however, any admission to touching the child's privates is a starting point to elicit a more serious admission when the suspect is brought into the police interrogation that will follow.

False Evidence

As reviewed earlier, police interrogators in the US can legally emphasize evidence against suspects, sometimes including false evidence, to further the belief that the suspects cannot convince others of their innocence and that they will be found guilty regardless of whether they confess. This feeling of hopelessness leads them to be more receptive to arguments that they will be viewed more favorably and obtain superior legal outcomes if they confess (Kassin et al., 2010). Pretext callers can also provide false evidence to make denial seem hopeless. For example, in the following exchange, acting on police instructions, the caller falsely claimed that witnesses saw the suspect sexually assault the caller.

    CALLER: "A and B (names omitted) saw. They were like, in the room. We've talked to them already. Like, everyone knows."

Until the presentation of this false evidence, the suspect claimed he would never commit such a crime, especially in front of family members, but he also found difficulty reconciling the apparent claims of multiple family members with his own memory. After the caller asserted the false evidence a second time along with a new name added to the list of witnesses, the suspect's defense clearly and immediately shifted from outright denial to failure of memory. This was the first in a series of small concessions, which eventually allowed the caller to take the suspect from "I never did to that to you; I would never do that" to "I'm sorry for touching you" after the caller continually demanded an apology for his alleged crimes (though he never admitted remembering having done it).

Limits on Pretext Callers

Pretext calls are not formal interrogations. Callers must balance efforts to gain evidence from the call with maintaining the secrecy of their true intent and law enforcement's involvement. This effort can prevent callers from

being as direct as police about what they want to hear. In addition, in many circumstances callers cannot realistically maintain a phone call as long as police interrogators can maintain a custodial interrogation, so callers must sometimes settle for weaker evidence than they want.

How Do Observers Interpret Suspect Statements to Pretext Callers?

One of the most ambiguous statements the suspect can make during a pretext call is an apology. Reasons for such apologies include admitting guilt, expressing remorse, soothing the caller without intending to admit culpability, or compliance to escape further demands.

One must also consider the nature of suspects' motives during a pretext call and whether such motives might prompt a false admission or a seemingly incriminating statement such as an apology. Many motives are likely, particularly in response to an accusatory pretext call. The accused is likely to want to prevent the caller from disclosing the accusations to police or others. He might also wish to manage the emotions of the caller or to help her deal with her distress (and make her feel better). He might simply wish to get the call over as quickly as possible without abruptly hanging up on the caller. Such motives might lead some suspects to comply with demands for an admission, apology, or explanation whether innocent or guilty, all the while unaware that their statements will be presented against them in court.

Lastly, unlike evidence related to many custodial interrogations, the jury has no access to video recording of the participants in the phone call. Without nonverbal cues, interpretation of the meaning of suspects' statements becomes more subject to error (Davis & Villalobos, 2014; Shuy, 1997). Listeners cannot see an expression of shock upon hearing an accusation, nor can they see facial expressions indicating irony or otherwise contradicting the literal content of the statements.

INFORMANTS

The deceptive practices law enforcement employs to obtain confessions are not strictly limited to getting a *direct* confession from the suspect, but might also include questionable methods to obtain what is known as a *secondary confession*. A secondary confession involves the testimony of a *cooperating witness* (Cassidy, 2004) who provides information about his or her knowledge of the suspect's involvement in a crime, obtained through interactions or conversations with a suspect (Neuschatz, Lawson, Swanner, Meissner, & Neuschatz, 2008). In other words, a cooperating witness would provide law enforcement with an account of a suspect's alleged confession, which could then be used as evidence in court. However, just like methods utilized to obtain primary confessions, those used to collect secondary confessions

have been criticized for relying on deception and contributing to wrongful convictions.

The Role of Informants in Covert Investigations

Informants provide investigative leads, facilitate casework, and may testify in court (Wilson, 1968). These functions are conducive to deceptive and ethically questionable practices by both law enforcement agents and informants themselves. There exists a lack of enforceable regulations regarding informant work (Natapoff, 2009). Generally speaking, the work of an informant consists of infiltrating or getting close to a target individual or group and obtaining leads or information that could be useful for a specific police investigation. Although the specific methods that informants use to gather information are varied and not extensively documented, the law enforcement agents to whom informants respond (often referred to as their *handlers*) expect informants to engage target suspects directly, spy, lie, eavesdrop, listen to conversations (occasionally while wiretapped), and do everything they can to obtain information undetected. Of particular interest, informants are usually tasked with obtaining some sort of admission or even a full confession for a crime, which could be used to justify an arrest or as evidence in court. Just as with pretext calls, suspects who confess to informants are not aware that by simply talking to someone they trust they might be inadvertently providing incriminating evidence against themselves to the police.

Miller (2011) made a clear distinction between what it means to be an *informer* vs. an *informant*, with the former being simply a person who would volunteer to *relay* information they already have to authorities, while the later would actively *seek* information "on behalf of authorities." Miller views this distinction as pivotal in understanding the questionable ethics of informant activity, as most of these individuals have criminal histories, and essentially function as "amateur officers" who work in exchange for rewards, many of which are illegal or morally ambiguous in nature. These rewards can range from money or drugs to promises of ignoring the informant's previous or current criminal activities, and even favorable discretion from prosecution in current or future criminal trials (US Department of Justice, 2005).

Informant operations are very diverse and can occur within the investigations of virtually any crime in which witness testimony can be used as evidence. Natapoff (2009) identified two types of snitching operations accounting for the majority: organized crime and drug-related. Informants working on organized crime operations are often monitored thoroughly, and their activities are extensively scrutinized, as these operations tend to be complex and long-lasting. Indeed, these informants might infiltrate crime organizations for months, or even years, before obtaining enough information to justify an arrest. In contrast, deals involving drug informants are much more common and involve fewer explicit guidelines. Such informants

can be virtually anyone willing to deceive the target: dealers, addicts, suspects' friends or family members, and even people who are already incarcerated or on probation. They can remain informants for any period of time and be asked to seek information about one or multiple people throughout the course of one or more investigations.

Law enforcement agents employ informants in a variety of settings. One prominent type of cooperating witness, the *jailhouse informant* (commonly known as a "snitch"), is used to gather information about an incarcerated suspect. Just like informants operating on the outside, jailhouse informants often agree or are pressured into seeking information from fellow inmates in exchange for ethically questionable rewards, such as monetary compensation, privileges or protections inside prison, or promises (explicit or otherwise) of a reduced sentence, or some other incentive for the defendant himself or someone else. Jailhouse informants operate in a similar way as their counterparts outside of prison, as their job is essentially to get close to a specific person or group and relay incriminatory statements or other evidence to their handlers in exchange for specific rewards.

Some such informants become highly skilled and are deployed repeatedly to elicit information from other inmates. Davis recently served in a case where, in a recorded interaction, a group of three repeat informants pretended to ally with the suspect, offered to help him obtain bail from friends outside prison, and systematically elicited information concerning the whereabouts of the murder weapon and other crucial evidence, and admissions consistent with first-degree murder, rather than self-defense.

(Lack of) Rules and Regulations in Informant Deals

Although informant deals vary widely, common themes do exist (Natapoff, 2009). Handlers tend to make first contact with potential informants, who can be targeted because of their informant potential or vulnerability (e.g., drug users in need of a way to sustain their habits; Miller, 2011), even if there is not enough evidence to warrant an arrest. Investigators are permitted to "bluff" or simply wait until the potential informant is involved in some sort of minor crime, such as possession of a controlled substance, to coerce cooperation. Regardless of circumstances, a deal is negotiated when the potential informant agrees to cooperate. Terms could be as general as a verbal promise not to charge the informant for a certain period of time or promising some sort of legal immunity, or as specific as full-on agreements detailing the informant's specific functions and rewards. At this point, the arrested suspect or targeted prisoner becomes an informant, and the recruiting officer becomes his handler. Each handler–informant relationship is different, and they might work on a single case or multiple cases, depending on the informant's access to criminal networks in or outside prison. Prosecutors might get involved later, as they have decision power to press formal charges and negotiate plea bargains in both ongoing and future trials.

The deals that informants and law enforcement make are notorious for being lopsided in favor of the informants, as informants are rarely prosecuted for lying or providing false information (Bloom, 2002). Informants are thus incentivized to provide any kind of incriminating information, truthful or otherwise, as they would have a lot to gain and virtually nothing to lose. There are often no requirements that information must have value, and as a result, many informants would purposely provide known information about the case to the police while presenting it as "findings" (Alter, 2005).

Prosecutors offering plea bargains to jailhouse informants often disregard the fact that these informants have little incentive to provide truthful information. Such situations pose ethical problems, as secondary confessions obtained through informants can add immense value to prosecutors' cases, especially when the remaining evidence for any given case is not particularly strong (Cassidy, 2004). The excessive flexibility in official discretion granted by the so-called informant law allows these ethically questionable practices to remain virtually unchecked (Natapoff, 2009).

Deception in Informant Testimony

When it comes to alleged informant-obtained confessions, the issue of informant deception is crucial. Secondary confessions often have more influence on jury decisions than other types of evidence (such as witness and character testimonies), and as much influence as primary confessions (Wetmore, Neuschatz, & Gronlund, 2014). Importantly, jurors might tend to believe informants are motivated to testify against a defendant for reasons such as feelings of guilt or sympathy for the family, and not to gain a reward or sentence reduction (Neuschatz et al., 2008). Importantly, Neuschatz and colleagues' (2008) study showed that explicitly telling jurors that the informant was offered a sentence reduction for testifying had no effect on sentencing decisions. Jurors appeared to accept the informant's testimony at face value and downplay the fact that the informant was given a very strong motive to lie.

The risk that untruthful secondary confessions might lead to wrongful convictions cannot be overstated. The Northwestern School of Law Center on Wrongful Convictions estimated that informants testified in the trials of 46% of wrongfully convicted individuals on death row who were exonerated between 1973 and 2004 (Warden, 2004). In addition, when looking at specific cases of exonerees, Garrett (2011) noted that 52 (21%) of the 250 cases he investigated involved informant testimony during trial. Likewise, The Innocence Project (2011, 2017) estimated that 20% of wrongful conviction cases overturned through DNA involved informant testimony.

Unfortunately, the safeguards established to protect defendants from faulty informant testimony—such as *Giglio v. United States* (1972) requiring prosecutors to disclose any kind of deals made with informants (Cassidy, 2004)—do not seem to be enough to prevent innocent people from being

incarcerated because of false alleged secondary confessions. For instance, the "reputation" of some prosecutors or investigators who have a history of reducing sentences of informants (Bloom, 2002) or an implied promise of help (e.g., "help us out and we'll take care of you"; Mazur, 2002) might be enough to get informants to testify, without the requirement of disclosing to the court that an "implied agreement" has taken place. If the prosecutor is successful in concealing the presence of a deal, jurors would not be aware of the informant's motives to lie and would not be able to examine the validity of the testimony on that basis.

Several parallels exist between how confessions are interpreted when obtained through an informant vs. a pretext call. In both instances, suspects are unaware that their statements could reach authorities. Both techniques rely on deceiving suspects into inadvertently admitting wrongdoings by trusting someone they might believe is "on their side." Additionally, confessions obtained through either means can be presented in court. In either case, statements made in-person or during recordings might be ambiguous, and prosecution, the pretext caller, or the informant can offer their own interpretation of any alleged admission. These interpretations are likely to favor their pre-existing beliefs and their own agendas. Importantly, as alluded to earlier, even if audio recordings exist, there is no way for jurors or anyone else to take into account the suspect's facial expressions, body language, and other idiosyncrasies that might modify the apparent meaning of a statement or conflict with the prosecution's interpretation of the alleged admissions.

In many ways, informant testimony is more misleading than statements obtained through police interrogations or pretext calls. Because informants might face high levels of risk to their well-being in or outside of prison if they are caught with any recording device, police and prosecutors are forced to rely solely on the word of the informant and his subjective interpretation of any statements made by the suspect. Most informants have strong incentives and a relative amount of freedom to omit, misinterpret, and outright lie about what they hear from and about the suspect, which makes their testimonies particularly dangerous, especially for innocent suspects.

Arguably, an even more difficult issue for interpretation of secondary confessions is that the lack of any recording or police presence makes it impossible to see what the informant did to elicit the confession. False confessions can occur for many reasons, including bragging or trying to make oneself seem tough to survive in prison. As is true for evaluating confessions obtained in any manner, one must know the context in which it occurred. Did the informant lie in ways that would encourage the suspect to claim credit just to seem important? Did he repeatedly state his conviction that the suspect was guilty until the suspect said, "Okay, you're right" to get the informant to shut up?

The use of informants to obtain evidence, especially in the form of secondary confessions, has elicited controversy among legal professionals, scholars, and the public at large. It is paramount that investigative and law agencies

enforce stricter rules and regulations to ensure that secondary confession evidence is reliable, such as holding pretrial hearings to verify the information or engaging in further investigation to corroborate the alleged confession (The Justice Project, 2007).

REFERENCES

Alter, V. (2005). Jailhouse informants: A lesson in e-snitching. *Journal of Technology Law & Policy, 10,* 223–243.

Bloom, R. M. (2002). *Ratting: The use and abuse of informants in the American justice system.* Westport, CT: Praeger Publishers.

Cassidy, R. M. (2004). Soft words of hope: Giglio, accomplice witnesses, and the problem of implied inducements. *North Western University Law Review, 98,* 1–43.

Cialdini, R. A. (2008). *Influence: Science and practice* (5th ed.). Upper Saddle River, NJ: Pearson.

Davis, D. (2008). Selling confession: The interrogator, the con man, and their weapons of influence. *Wisconsin Defender, 16*(1), 1–16.

Davis, D. (2010). Lies, damned lies, and the path from police interrogation to wrongful conviction. In M. H. Gonzales, C. Tavris, & J. Aronson (Eds.), *The scientist and the humanist: A festschrift in honor of Elliot Aronson* (pp. 211–247). New York, NY, USA: Psychology Press.

Davis, D., & Leo, R. A. (2014). The problem of interrogation-induced false confession: Sources of failure in prevention and detection. In S. Morewitz & M. Goldstein (Eds.), *Handbook of forensic sociology and psychology* (pp. 47–75). New York: Springer.

Davis, D., Leo, R. A., & Follette, W. C. (2010). Selling confession: Setting the stage with the "sympathetic detective with a time-limited offer". *Journal of Contemporary Criminal Justice, 26*(4), 441–457.

Davis, D., & O'Donohue, W. T. (2004). The road to perdition: Extreme influence tactics in the interrogation room. In W. T. O'Donohue & E. Levensky (Eds.), *Handbook of forensic psychology* (pp. 897–996). New York, NY: Elsevier Academic Press.

Davis, D., & Villalobos, J. G. (2014). Language and the law: Illustrations from cases of disputed sexual consent. In T. M. Holtgraves (Ed.), *The Oxford handbook of language and social psychology* (pp. 438–458). New York, NY, USA: Oxford University Press.

Drizin, S. A., & Leo, R. A. (2004). The problem of false confessions in the post-DNA world. *North Carolina Law Review, 82,* 891–1007.

Garrett, B. L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong.* Cambridge, MA, USA: Harvard University Press.

Giglio v. United States, 405 US 150. (1972). 405 US 150.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2013). *Criminal interrogation and confessions* (5th ed.). Burlington, MA: Jones & Bartlett Learning.

Justice Project. (2007). *Jailhouse snitch testimony: A policy review*. Washington, DC: The Justice Project.

Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist, 52*(3), 221–233.

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior, 34,* 3–38.

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A. M., et al. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior, 31*(4), 381–400.

Lavine, E. (1930). *The third degree: A detailed and appalling expose of police brutality.* New York, NY: Garden City.

Lavine, E. (1936). *Secrets of the metropolitan police.* New York: Garden City.

Leo, R. A. (1996). Inside the interrogation room. *The Journal of Criminal Law and Criminology, 86*(2), 266–303.

Leo, R. A. (2008). *Police interrogation and American justice.* Cambridge, MA: Harvard University Press.

Leo, R. A., & Davis, D. (2010). From false confession to wrongful conviction: Seven psychological processes. *The Journal of Psychiatry & Law, 38,* 10–56.

Mazur, E. P. (2002). Rational expectations of leniency: Implicit plea agreements and the prosecutor's role as a minister of justice. *Duke Law Journal, 51,* 1333.

Meissner, C. A., Redlich, A. R., Michael, S. W., Evans, J. R., Camilletti, C. R., Bhatt, S., et al. (2014). Accusatorial and information-gathering interrogation methods and their effects on true and false confessions: A meta-analytic review. *Journal of Experimental Criminology, 10*(4), 459–486. https://doi.org/10.1007/s11292-014-9207-6.

Miller, J. M. (2011). Becoming an informant. *Justice Quarterly, 28*(2), 203–220.

Miranda v. Arizona, 384 U.S. 336. (1966).

Natapoff, A. (2009). *Snitching: Criminal informants and the erosion of American justice.* New York: New York University Press.

Neuschatz, J. S., Lawson, D. S., Swanner, J. K., Meissner, C. A., & Neuschatz, J. S. (2008). The effects of accomplice witnesses and jailhouse informants on jury decision making. *Law and Human Behavior, 32*(2), 137–149. https://doi.org/10.1007/s10979-007-9100-1.

Scherr, K. C., & Madon, S. (2013). "Go ahead and sign": An experimental examination of Miranda waivers and comprehension. *Law and Human Behavior, 37*(3), 208–218.

Shuy, R. W. (1997). *The language of confession, interrogation, and deception* (Vol. 2). Thousand Oaks, CA: Sage.

The Innocence Project [innocenceproject]. (2011, July 15). *Getting it right: Informants.* Retrieved from https://www.youtube.com/watch?v=MM700n64Bx8.

The Innocence Project. (2017). *DNA exonerations in the United States.* Retrieved June 20, 2017 from https://www.innocenceproject.org/dna-exonerations-in-the-united-states/.

Thomas, G. C. I. (1996). Plain talk about the Miranda empirical debate: A "steady-state" theory of confessions. *UCLA Law Review, 43,* 933–959.

U.S. Department of Justice. (2005). *The Drug enforcement administration's payments to confidential sources.* Washington, DC: U.S. Department of Justice.

Villalobos, J. G., Kemmelmeier, M., & Zimmerman, C. A. (2018). Trust in the law: Why authoritarians are more likely to recommend falsely confessing. Manuscript submitted for publication.

Warden, R. (2004). The snitch system: How incentivized witnesses put 38 innocent Americans on death row. *Northwestern University School of Law*. Center on Wrongful Convictions, 1–16.

Wetmore, S. A., Neuschatz, J. S., & Gronlund, S. D. (2014). On the power of secondary confession evidence. *Psychology, Crime & Law, 20*(4), 339–357. https://doi.org/10.1080/1068316X.2013.777963.

Wicklander-Zulawski & Associates, Inc. (2017). *Wicklander-Zulawski discontinues Reid Method instruction after more than 30 years* [Press release]. Retrieved from https://www.w-z.com/portfolio_category/wz-news/.

Wilson, J. Q. (1968). *Varieties of police behavior*. Cambridge, MA: Harvard University Press.

Interrogation and Torture: The Dark Side of Deception and Law Enforcement

*Daniel Cochece Davis, Cynthia Adarian Wartanian,
Kimberly Beach and Danielle Blake Prentice*

When deception occurs in law enforcement contexts, most people assume it pertains to the suspect engaging in deceptive behavior. While this is often the case, deception is also used as a law enforcement interrogation technique. While deception during interrogation interviews could be considered unethical, dire public safety and national security conditions may provide justification to many of its uses. Yet, tactics used in the interrogation process can be problematic, inhumane, and ineffective. This chapter focuses on "dark side" issues (Spitzberg & Cupach, 1994) pertaining

D. C. Davis (✉)
School of Communication, Illinois State University, Normal, IL, USA
e-mail: dadavis@ilstu.edu

C. A. Wartanian
Long Island, NY, USA

K. Beach
Torrington, CT, USA

D. B. Prentice
San Diego, CA, USA

to interrogation techniques used by law enforcement personnel to uncover perceived deception and search for the ground truth. It is an effort to review, acknowledge, and further unearth interrogation's dark side, as well as call into question the continued use of problematic interrogation techniques. Further, this review focuses on interrogation practices, distinguishing their legitimate application from torture. We also focus on the communicative aspects within criminal and terrorist interviews, highlighting interrogation techniques and false confessions.

While many aspects of torture and interrogation are consistent across history—motivating reluctant individuals to disclose information, interrogators' goals of discerning truthful from deceptive information—this topic is especially relevant, as the 45th President of the United States, Donald Trump, was elected on a campaign supporting extreme interrogation techniques bordering on, if not including, torture, as "effective" and useful (Stableford, 2018). Following the election, he nominated Gina Haspel for Director of the Central Intelligence Agency (CIA), who supervised the CIA's program of torturing terrorist suspects (Papenfuss, 2018; Stableford, 2018), as well as facilitated the destruction of videos recording these events (Riechmann, 2018).

The use of extreme interrogation and torture techniques may seem unethical, if not illegal. Yet, torture was actually historically *required* by some courts, with confessions "only legally admissible if elicited under torture because other testimonies were deemed less reliable and valid" (Davis, 2014a, p. 893). This perspective assumed individuals would be honest under such harsh conditions. Over time, courts have reversed their position, believing that tortured people will say anything to stop the interrogation process (Rejali, 2007). The fictional character Portia, in William Shakespeare's *The Merchant of Venice*, captures this aspect well when he says "Ay, but I fear you speak upon the rack. Where men enforcèd do speak anything" (1995, Act 3, Scene 2, p. 2). Thus, if the purpose of law enforcement interviews is to gain valid information, the techniques used must be reliable in gaining truthful suspect information.

In the current chapter, we review the legal components and constraints of interrogations within the context of the US. Following this, we provide a review of common interrogation tactics, including the communicative dimensions, components, and outcomes of coercive behaviors such as torture, sensory deprivation, and strategic use of interrogators' nonverbal behavior. Collectively, this analysis reveals that deception cannot be viewed simply as an ethical or unethical communicative endeavor, but instead has the capacity for being viewed as either, or somewhere in between, depending on its social and cultural framing, as well as the net yield of related safety and security outcomes.

Law Enforcement Constraints: Miranda Rights and Relevant Constitutional Amendments

In order to study the deceptive elements of interrogation, it is important to understand applicable legal components that constrain law enforcement activities. The Miranda Rights, mandated in 1966 in *Miranda v. Arizona* (Shuy, 1998), are read aloud and administered by law enforcement officials prior to a custodial interrogation as a way of explaining to arrested persons their legal rights (Zulawski & Wicklander, 1992). The Miranda Rights for suspects are: (a) the suspect has the right to remain silent; (b) anything the suspect says can and will be used against them in a court of law; (c) the suspect has the right to talk to an attorney and have them present while being questioned; (d) if the suspect cannot afford to hire an attorney, one will be appointed to represent them before any questioning, if they wish one; and (e) the suspect can decide at any time to exercise these rights and not answer any questions or make any statements. A pre-interrogation waiver verifies that the suspect comprehends these rights (i.e., by asking if the suspect understands each of these rights explained to them). Having these rights in mind, the suspect is then asked if they wish to talk (Graham, 1970). These rights are foundationally based on amendments to the original US Constitution. The Fifth Amendment to the US Constitution protects persons from testifying against themselves by providing the right to remain silent (Middleton, Lee, & Chamberlin, 2004). The Sixth Amendment provides any person accused of a crime the right to counsel (Zulawski & Wicklander, 1992). Those who are accused are also granted the right to a trial by an impartial jury—a group of persons whose knowledge of the trial is limited to the courtroom's boundaries (Middleton et al., 2004).

These amendments aim in part to ensure a just courtroom trial; however, these constitutional amendments fail to focus on obtaining essential information or evidence in a reliable manner during investigations (Dripps, 1999). Common law has its own voluntariness test to determine whether a confession is voluntary, and the confession evidence is to be discarded if found untrustworthy (Dripps, 2003; Graham, 1970; White, 1998). Nagel (1972) states the Fifth Amendment is also responsible for deciding if a confession is to be considered incompetent due to its level of voluntariness. As such, evidence being collected has very few constraints. Similarly, a portion of the Fourteenth Amendment can be seen as an elaboration of the Fifth Amendment, explaining that, "no state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any state deprive any person of life, liberty, or property, without due process of law, nor deny to any person within its jurisdiction the equal protection of the laws" (Zulawski & Wicklander, 1992, p. 33). This amendment protects persons being interrogated from producing either an involuntary or

untrustworthy confession (Grano, 1993), resulting from the use or threat of physical force or other coercive tactics (Caplan, 1985; Middleton et al., 2004; White, 2001b). While these protections provide a type of broad coverage to all citizens, interrogators working within these boundaries still have significant latitude in using communicative techniques to extract information and ascertain the veracity of information that has been provided. To-date, several authors provide widely accepted and used accounts of the interrogation process (e.g., Davis, Leslie, & Davis, 2014; Hess, 2015; Inbau, Reid, Buckley, & Jayne, 2011; O'Hara & O'Hara, 1998). Just as individuals tailor their messages specifically to receivers in order to communicate effectively (Trenholm, 2005), skilled interrogators will carefully use relevant techniques that fit within the given legal constraints and to the specific types of suspects, customizing the interrogation instead of applying a general interrogation script to all situations.

The Interrogation Process and Considerations

A common belief guiding many criminal investigators is that the suspect's confession is the most influential part of a case. Many investigators feel a case cannot be won without the suspect's confession (Kassin, 1997; Magid, 2001), and once the confession is obtained, all other aspects of the trial become superfluous (Kassin, 1997). In order to obtain a confession, interrogation techniques are designed to find each suspect's weakness and exploit it (White, 1997). The goal of interrogations, which are predominately *accusatory* in nature, is to gain compliance (Cody, Canary, & Smith, 1994) in hopes of obtaining the truth (Gordon & Fleisher, 2002; Inbau et al., 2011). Effective interrogators gain compliance by sometimes deceptively decreasing suspects' perceptions of the consequences of confessing, while simultaneously increasing the suspect's internal anxiety about their own deception (Grano, 1986). If the interrogator is successful in increasing suspects' anxiety, suspects will believe that continuing to deny their guilt will lead to even increased anxiety, while admitting to their guilt will eliminate it (White, 1997).

Though a confession is believed to be the most influential part of a case, in reality the importance of a confession depends largely on the strength of the facts in the case, as well as how effective this evidence alone can lead to a conviction (Gudjonsson, 1992). Evidence may include witness statements, which can be based on faulty human memory leading to "close to three-quarters of wrongful convictions" (Geddes, 2016, p. 37). Indeed, many false, deceptive, or coerced witness statements are a reason for false indictment, though these are often dismissed by courts before convictions (Geddes, 2016; Hail-Jares, Lowrey-Kinberg, Dunn, & Gould, 2017). Therefore, it is necessary for interrogators to have consistent measures to judge if the *witness*, in addition to the suspect, is telling the truth (Strömwall & Granhag, 2003). Geddes (2016) outlines measures to help preserve witness accuracy, including:

limiting exposure of the witness to a suspect before formal identification takes place; blinding the person administering the line-up to who the real suspect is; telling the witness that the perpetrator may or may not be in the line-up, so they don't feel pressured to choose someone; and recording their confidence in the identification, should they make one. (p. 38)

The challenge for general law enforcement officials, though, is that the accuracy rate of distinguishing a liar from a truth teller is below 60%, being only slightly higher than chance (Granhag & Strömwall, 2001a). Even more to the point, Geddes (2016) notes that "studies have revealed that even trained police interrogators only get it right just over half the time—about the same as inexperienced college students" (p. 38). Thus, a common, yet false assumption is that those trained in interrogation are more effective at detecting deception than both the general public and those law enforcement individuals not trained in interrogation.

Due to the perceived significance of confessions, police officers may use every possible tool to obtain confessions from those suspects they are investigating (Helm, 2003). Often, the only witnesses to a crime are the criminals themselves; therefore, obtaining their testimony and confession is imperative (Maver, 1996). While a general belief held by both law enforcement and the general public is that interrogation practices are productive, interrogators should not persuade innocent persons to confess to crimes they did not commit (Grano, 1986).

As stated, the method of determining a truthful statement has varied over time, with some interrogators historically *legally bound* to extract statements under torture to ensure the information was not deceptive. In current interrogation contexts, there are several subtle aspects of the suspect's environment and an interrogator's techniques that work to facilitate the interrogation.

Interrogation Techniques and the Use of Deception

Skilled interrogators often use verbal and psychological interrogation techniques, rather than physical ones (e.g., Alison, Giles, & McGuire, 2015; Dawes, 2004). Indeed, some interrogation techniques include deception. For example, police officers may encourage a confession by deceptively misrepresenting a case's facts to suspects, who might then have reason to confess if the facts provided were true (Sasaki, 1988). Whether interrogators fabricate evidence (Young, 1996) or make false claims about the status of "incriminating" evidence, suspects are often deceptively persuaded to believe there is actually a stronger case against them (Aubry & Caputo, 1980; Wakefield & Underwager, 1998; Woody & Forrest, 2009; Woody, Forrest, & Stewart, 2011).

Manipulation of Language and Information

In attempts to gain a confession, suspects may be told that an accomplice has already confessed or has provided sufficient evidence for a conviction (Slobogin, 1997), or that there are eyewitnesses who can identify the suspect (Magid, 2001). Officers sometimes offer contradictory or incorrect information about the case, hoping suspects will clarify the fact, thereby catching themselves in a confession (Blagrove, 1996; Inbau et al., 2011). Suggesting jail terms or alluding to the seriousness of the suspect's sentencing, which may be reduced or increased according to cooperation, is another example of misrepresented facts (Wakefield & Underwager, 1998). Suspects can be fairly easily led to believe that it would be in their best interest to confess (Magid, 2001).

Police might use specific deceptive wording (Buller & Burgoon, 1994) and behavioral techniques that take advantage of a suspect's emotions. These persuasive emotional appeals are another interrogation technique used to deceive suspects (Sasaki, 1988). Officers may pretend to befriend suspects or give false impressions of sympathy for their situation. They may suggest excuses or moral justifications for the crime, suggest the victim was to blame, or appeal to the suspect's religious beliefs (Kassin, 1997). Some officers have been known to pray with their suspects, hoping to create a situation where the suspect feels emotionally comfortable confessing to the crime under investigation (Young, 1996).

Interrogating officers may suggest to suspects that they are causing family and friends harm through the situational creation of stress (White, 2001a), or forbid contact with family or lawyers (International Injustice, 2004). Isolation and confinement can cause a variety of psychological and behavioral disturbances and may result in a suspect's loss of contact with reality. This has the possibility of leading to an eventual confession (Wakefield & Underwager, 1998). Scare tactics (Kassin, 1997), as well as alternating rewards and punishments to the suspect, are other ways to psychologically alter a suspect's willingness to confess (Solomon, 2003).

Another deceptive technique practiced by some interrogators is the intentional failure to inform the suspect of an influential fact of the case (Sasaki, 1988). The knowledge of this fact may influence whether the suspect does or does not confess. The "two-step method" is a form of this deceptive technique. With this method, a confession or incriminating statement is first drawn out of the suspect, and then, the Miranda Rights are read. Suspects are under the impression that they have already confessed, and believe repeating their story will not have any further harm. What suspects do not realize is that the initial confession cannot be used in the case against them, and unless the confession is repeated *after* the Miranda Rights are read, they have not incriminated themselves (Helm, 2003). Suspects may choose to enact their Miranda Rights at any time, and the police are then required to end questioning (Cassell, 1998; Cassell & Hayman, 1996).

Promises are another interrogation technique frequently used to extract confessions. Promises of leniency are occasionally effective in obtaining a confession, particularly in lengthy interrogations (White, 1997). Often, suspects feel as though they owe something to interrogators who promise leniency, even though such leniency can only be granted by a judge. This technique is referred to as *reciprocity* (Cialdini, 1998), meaning it is expected for individuals to return favors (Seiter & Gass, 2004). Thus, a suspect promised leniency might confess because doing so would please the interrogator (White, 1997).

As mentioned, interrogators may falsely claim they have evidence against the suspect, such as fingerprints, hair, blood, semen, or witnesses (Shepard, 1991). This deception is an intentional act in which senders knowingly transmit messages intended to foster a false belief or interpretation by the receiver (Seiter & Gass, 2004). During an interview, suspects often have distorted views of reality regarding information and its handling, as well as the relationship between themself, the interviewer, and the outside world. This commonly leads to suspects making choices that they would not usually make (Shepard, 1991). The purpose of fabricated evidence is to scare suspects who resist confessing and to cull out additional information (Alschuler, 1997).

Increasing Suspect Suggestibility and Anxiety

When interrogators think someone may be lying, they often interrogate for long periods of time, or more than once, in order to detect deception via inconsistencies across the repeated interrogations (Granhag & Strömwall, 2001a, 2001b, 2002). Suspects suffering from sleep deprivation, a result of a lengthy interrogation, have a difficult time differentiating between the interrogator's regular and misleading information. In these situations, investigators ask leading questions because sleep-deprived individuals have increased levels of *suggestibility* (Blagrove, 1996; Gudjonsson, 2003). Similarly, by interrogating at night and interrupting sleep patterns or eliminating time for sleep, truthful confessions will more easily be obtained from overtired, stressed, guilty suspects (Blagrove, 1996; Davis, 2014a).

Along with sleep deprivation, there are many other aspects of the interrogation process causing suspects' stress and increasing their anxiety levels or arousal (Davis, 2014b). A suspect's inherent *arousal level* can affect how anxious they are while being interrogated (Buller & Burgoon, 1994; Davis, 2014b; O'Hair & Cody, 1994). A person with high arousal levels could quickly become overstimulated during an interrogation and, therefore, may confess more easily in order to end an interrogation and return to a more comfortable arousal level (Zuckerman, 1994; Zulawski & Wicklander, 1992). Conversely, persons with low arousal levels would be better able to withstand lengthy interrogations, requiring more intensive interrogation techniques to be used in order to reach overstimulated arousal levels (Libkuman, Griffith, Wines, Dickel, & Doty, 1998).

The way in which the interrogator speaks to the suspect or witness can also increase arousal levels. If interrogators yell or vary their vocal patterns, the person being interrogated may become more stimulated, potentially leading to higher anxiety. In addition, worrying over whether they are assumed to be guilty, what will happen to them during the investigation, as well as what the police will uncover about their past, all increase anxiety. As with other interrogation techniques, the success in getting a guilty suspect to confess is offset with innocent suspects becoming more likely to falsely confess (Stewart, Woody, & Pulos, 2018). This dark side or perhaps "dual-sided" aspect of interrogation techniques must be acknowledged and monitored for false confessions. Doing so may very well help avoid false confessions and more broadly aid in advancing just practices.

Nonverbal Contributors to Deception Detection

The physical environment of the interrogation can also cause stress and higher anxiety in suspects (Gudjonsson, 2003). The term *proxemics* refers to the nonverbal use of space and can be highly influential in persuasive situations such as interrogations (Seiter & Gass, 2004; Yeschke, 2003). Though not thoroughly researched by deception scholars, many deception detection practitioners understand that the room in which the interrogation takes place—as well as the interrogator's use of personal space—potentially affects interactions during the interrogation and its outcome. Practitioner sources indicate that bare, small, and soundproof rooms with only simple chairs and a desk are ideal for interrogations because they are the least distracting (Aubry & Caputo, 1980; Macdonald & Michaud, 1987). Rooms without phones or any switches within reach are also effective at keeping the interview focused (Dowling, 1979; Kassin, 1997). Rooms without furniture or other proxemic barriers between the suspect and the interrogator help facilitate conversation (Grano, 1986; Walkley, 1987). The interrogator's chair is often higher than the suspect's chair to give a perception of superiority. Further, various sources indicate that interrogators' chairs on casters allows them to move into suspects' spaces as desired, prompting suspects to confess out of proxemic discomfort (Gordon & Fleisher, 2002; Yeschke, 2003). Related, any tension-relieving items or actions may be banned from the room, such as cigarette smoking or small objects with which to play. By increasing the suspect's tension or, conversely, not allowing for a tension decrease, the suspect may be more likely to confess (Grano, 1986). Though these environmental elements can facilitate a guilty suspect's confession, they may also facilitate greater probability of an innocent suspect providing a false confession.

In addition to the physical environment, individuals' nonverbal behavior and physiological cues are also relevant elements pertaining to deception

detection in the contexts of interrogations. Since it is difficult to consciously control all behavioral channels (i.e., body movements, eye behaviors, vocal behavior), nonverbal communication that conflicts with verbal statements often "leaks out" through facial, vocal, and/or bodily expressions (Ambady & Rosenthal, 1992). This creates a drop in message fidelity, potentially indicating a suspect's deceptive attempt. Davis (2014b) remarks, "With so many channels to simultaneously attend to, deceivers would often 'leak' behaviors associated with the truth, mixed in with controlled behaviors associated with their lie, creating a disparity across the behavioral channels" (p. 704). What a person is thinking affects how they behave, and interrogators have established ways of attempting to read these cues.

One way law enforcement interrogations and interviews use nonverbal and physiological cues concerns the use of the polygraph. While polygraph technology monitors suspects' heart rate, blood pressure, and the skin's electrical conductance via varying levels of water being present (Davis, 2014c), polygraph administrators seek suspects' verbal statements (i.e., confessions) and use nonverbal and physiological cues as indicators of where to continue probing for information (Davis, 2014c; Horvath, 1973). Thus, "a goal of a polygraph interview is to have participants make verbal statements that reveal guilt, if appropriate. Courts would admit such verbal statements as confessional evidence" (Davis, 2014c, p. 775). As with any confession, courts are primarily concerned with the validity of the confession. While polygraphs can facilitate accurate statements by suspects, other interrogation techniques, some of which have been mentioned, can lead to false confessions.

INTERROGATION PRACTICES AND FALSE CONFESSIONS

A false confession occurs when a suspect admits to being guilty of a crime or crimes that they did not commit (Gudjonsson, 2003). Suspects often confess, despite the fact they are innocent, in hopes of gaining notoriety, shelter (Davis, 2014a; Macdonald & Michaud, 1987), or, controversially, because they are pressured by interrogators (Davis, 2014a; Leo & Ofshe, 1998). The use of coercive methods, creating extreme anxiety, interrogating those individuals who are most susceptible, and ignoring measures that can help prevent false confessions all increase the likelihood that a false confession will occur.

Coercion and False Confessions

Coercive methods used in interrogations can cause innocent suspects to claim guilt. By definition, coercion is "any irresistible or overwhelming inducement" (Alschuler, 1997, p. 960). The Supreme Court states that coercion occurs whenever the defendant's will is overborne (Alschuler, 1997). Alschuler suggests that coerced confessions are caused by offensive

governmental conduct such as deception, promises, and threats. As such, suspect confessions obtained via interrogators' use of deception could be thrown out as evidence, diminishing the probability of successful prosecution. Police asserted in *Miranda v. Arizona* that interrogation is in accordance with the Constitution, "provided the tactics employed were not so coercive as to induce an innocent person to confess" (Skolnick & Fyfe, 1993, p. 57). Thus, by definition, any false confession, resulting from interrogation pressures, is a "coerced" confession.

The use of coercion in acquiring confessions is founded on an aspect of persuasive communication where the coercive person engages specifically in "compliance-gaining" tactics to influence cooperation (Spitzberg, Marshall, & Cupach, 2001), or "a type of planned, goal-directed communication" (Seiter & Gass, 2004, p. 20). In interrogations, there are two types of coerced false confessions: coerced-compliant and coerced-internalized. Coerced-*compliant* false confessions are when suspects confess in order to obtain a goal, or to relieve themselves from a difficult position or situation (Gudjonsson, 1992; White, 1997). The coercive pressures of the interrogation process lead to suspects' confessions when they involuntarily succumb to the persuasive pressures and demands of the interrogator for an immediate gain (Gudjonsson, 1992). This communicative action is perpetuated by the interrogator's intent and her or his manipulation of cues and messages to induce certain responses and confessions (O'Hair & Cody, 1994).

The second type of coerced confession is the coerced-*internalized* false confession, which occurs when suspects become uncertain of their own innocence (Gudjonsson, 1992). Suspects often become confused and begin to doubt themselves because of manipulative and coercive interrogation tactics adjusting their perceptions of reality (Buller & Burgoon, 1994; Gudjonsson, 1992). One notable example of coerced-internalized false confessions is when 18-year-old Peter Reilly confessed to killing his mother after ten hours of interrogation (White, 1997). Reilly began to believe that he had committed the accused crime during the interview (Gudjonsson, 1992). Thus, while coerced confessions may appear to result in usable evidence, too often they emerge as false, squandering valuable law enforcement time and resources, while also delaying capture of actual perpetrators.

Similarly, suspects who are not guilty often still feel anxious, potentially leading to false confessions. This anxiety sometimes occurs because the suspect has lied to the interrogator about an issue unrelated to the crime. For example, a suspect might lie about his whereabouts when his wife was murdered, because at the time he was in bed with another female (Connery, 1977). Condit (1997) and Lunbeck (1994) suggest that females tend to be more secretive about their personal lives and might therefore lie in order to avoid unwanted questioning. Suspects can experience uneasiness or increased anxiety when arguing with strong authority figures, such as police officers and prosecutors (Connery, 1977), especially if they perceive that these authority

figures will not accept or consider the suspect's position (White, 1997). Conversely, the suspect's anxiety from the pressures of interrogators may be lessened by the presence of legal counsel (Caplan, 1985), thereby decreasing the possibilities of a false confession.

Those Most Susceptible to Falsely Confessing

Individuals who have intellectual disabilities are the most prone to falsely confessing to a crime. In one study, Leo and Ofshe (1998, as cited in White, 2001b) found that of sixty proven or probable false confessions, at least seventeen were by suspects who had intellectual disabilities. These individuals have a tendency to accept and be persuaded by what the police tell them, as many are extremely susceptible to coercion and pressure. They may also desire to please the interrogator who plays an authoritative role (White, 1997). Unfortunately, due to the lack of observable signs, it is often difficult to identify individuals with intellectual disabilities prior to, or during interrogations (Gudjonsson, Clare, Rutter, & Pearse, 1993). Therefore, interrogators may unknowingly receive false confessions from these suspects.

Juveniles and those with compliant personalities are also more likely to falsely confess (Macdonald & Michaud, 1987). Gudjonsson (2003) characterizes the compliant personality as having two major traits. Those with this personality type have an eagerness to please and protect their self-esteem when in the company of others, as well as tendencies to avoid conflict and confrontation, in particular with authoritative figures (White, 1997). Feelings of guilt, discomfort discussing unfamiliar or sensitive topics, or undeveloped communication skills may cause a child to unknowingly give false information (Inbau et al., 2011). Due to these traits, these suspect types are especially vulnerable to law enforcement interrogation methods, and interrogators should be vigilant to the possibilities of obtaining false confessions from them.

Measures to Prevent False Confessions

There are numerous measures to prevent false confessions. In particular, experts agree that the length of time in the interrogation should be limited (Blagrove, 1996; Cassell, 1999; Gordon & Fleisher, 2002; Hancock, 2003). "Some courts have allowed interrogations longer than five hours; however, lengthy interrogations signify to the suspect that the only way to end the interrogation is to give in and confess" (Davis, 2014a, p. 895; see also Alschuler, 1997). Once an interrogation reaches a certain length, a suspect's ability to protect him- or herself against an interrogator's pressure deteriorates. The intensity of interrogations conducted on juveniles and individuals with intellectual disabilities should also be limited (Cleary & Warner, 2017; Eastwood, Snook, Luther, & Freedman, 2016; Freedman, Eastwood, Snook, & Luther, 2014; Liefaard & van den Brink, 2014). Inbau, Reid, and Buckley

(1986) claim, "Special protection must be afforded to persons of below average intelligence…to minimize the risk of obtaining untruthful admissions due to their vulnerability to suggestiveness with respect to possible explanatory conduct" (p. 195).

Another potentially effective measure in preventing false confessions is videotaping or recording interrogations (Hancock, 2003; Shuy, 1998; White, 2001b). As of 2003, one-third of all US law enforcement agencies already took this measure (Hancock, 2003). England requires all police interrogations to be videotaped (Alschuler, 1997). An English Royal Commission stated, "By general consent, tape recording in the police station has proved to be a strikingly successful innovation providing safeguards for the suspect and the police officer alike" (Alschuler, 1997, p. 977). Videotaping does not jeopardize legitimate law enforcement agencies (Davis, 2014a; White, 1997), and "when interrogations are fair, it produces powerful evidence for prosecution" (Davis, 2014a, p. 895; see also Alschuler, 1997; Cassell, 1999).

The Interrogation of Terrorists

In addition to more typical criminal investigations, interrogations are also vital in helping to fight terrorism. However, the methods used to interrogate terrorists and suspected terrorists are not always humane. Thus, while the outcomes may be positive (e.g., disclosure of information that prevents attacks or assists missions), the means by which they are achieved can be quite negative (Parry & White, 2002; Spitzberg & Cupach, 1994) and potentially result in both short- and long-term consequences (e.g., injury, psychological damage, false confessions) (Ries, 2017). To guard against such potentially damaging results, countries can agree to an international code of prisoner ethics referred to as the Geneva Conventions. Article 3, which is found in all four Geneva Conventions, is an effort to protect prisoners of war by setting standards for humane treatment for detainees, under all conceivable circumstances (Paine, 2004). While interrogations may be used to determine whether those suspected of terrorism are deceiving their captors, techniques considered "torture" violate the Geneva Conventions.

Torture

Interrogation and torture are inherently communicative activities (Davis, 2014a). The most controversial interrogation technique dealing with terrorism is the use of torture (Mackey, 2004). Title 18, Section 2340 of the US Code defines torture as "an act committed by a person acting under the color of law specifically intended to inflict severe physical or mental pain or suffering (other than pain or suffering incidental to lawful sanctions) upon another person within his custody or physical control" (Massimino, 2004, pp. 74–75; see also York, 2004). Graessner, Gurris, and Pross (2001) state that the "goal

of torture is to destroy the personality and annihilate identity. Torturers know that people without identity—people with shattered personalities—lose their capacity for resistance and give in to the demand that they reveal secrets and practice betrayal" (p. xi). Davis (2014a) notes that Bravin and Fields (2003) argue that, in the US, "as long as the pain and suffering inflicted on a prisoner is not 'severe,' it is permissible to use physical force to cause 'discomfort'" (p. 893). Yet the realities and awareness of torture and terrorism, through such instances of coercion and dominance in quests to expose deception and truth, also expose a severely dark side of communication due to resulting physical and psychological abuse (Marshall, 1994; Spitzberg & Cupach, 1994). While interrogation within the US is constrained by legal parameters such as Miranda Rights and constitutional civil liberties, methods to subvert or work around these constraints facilitate the problematic dark side of interrogation.

The United States and Terrorism

Reports on the misuse of interrogation techniques when dealing with terrorists suspected of threatening US national security raise the question of whether limits on interrogation should be more strictly enforced (Parry & White, 2002). The US constitutional limits on police interrogators are dependent upon the interrogation's purpose. If there is an immediate concern for protecting the public, such as terrorist actions, the Supreme Court has held that agents may question the suspected terrorists without first warning them of their Miranda Rights (Parry & White, 2002). This is consistent with Massimino's (2004) assertion that under the George W. Bush administration, there was "a fundamental shift in approach when dealing with terrorist suspects, from prosecution to prevention" (p. 74). Accompanying this shift was an abandonment of Miranda Rights, appearing in court, and having legal representation present, as well as other typical legal protocols.

Various reports claim that the US has tortured terrorist suspects or is involved in tactics similar to torture (Balfe, 2018; Jones & Sheets, 2009; Smith, Shane, Mazzetti, & Baker, 2015). The US has interrogated terrorist suspects using torturous methods such as prolonged isolation in bare, dark cells, exposure to heat and cold, sexual humiliation, and diet manipulation (Paine, 2004). The US also uses stress and duress methods (CIA, 2002; Smith et al., 2015) in the interrogation of terrorists, such as 24-hour bombardments of light, hooding, sleep deprivation, and the forcing of prisoners to hold awkward positions for hours on end (Bowden, 2003; Ends, 2003; Finn & Warrick, 2009; Hentoff, 2003; Panetta & Serraglio, 2000; Stein, 2014). Due to US laws limiting even harsher interrogation techniques, those terrorist suspects viewed as the most "hardened" have been relocated (i.e., "extraordinary rendition," "irregular rendition," or "forced rendition") to other countries specifically known for their brutality in torturing prisoners, such as Jordan, Morocco, Syria, and Egypt (Bravin & Fields, 2003; Howe,

2003; Smith et al., 2015). This is in direct violation of the United Nations Convention Against Torture (Davis, 2014a).

US military interrogators of 9/11 terrorist suspects used techniques including deception, screaming, and the use of false evidence. Prisoners were stripped, forcibly shaved, deprived of religious items and toiletries, and medical personnel assisted by using rectal hydration and other interventions (Balfe, 2018), even though these might be medically unnecessary or even potentially life-threatening via water intoxication if used in conjunction with waterboarding techniques ("Rectal Hydration," 2014). Military interrogators also play on prisoners' fears, such as placing rats or dogs near or on them (Bravin & Fields, 2003). One official who had supervised the capture and transfer of accused terrorists was quoted as saying, "If you don't violate someone's human rights some of the time, you probably aren't doing your job" (CIA, 2002, p. 30). However, many suspects are trained in the very same extreme interrogation methods in order to resist law enforcement interrogation, reducing the effectiveness of these methods (MacDonald & Michaud, 1987). In the US, many scholars and skilled interrogators have concluded that these harsh physical methods are often "less effective than psychologically-oriented techniques in obtaining truthful statements from terrorists" (Davis, 2014a, p. 893; see also Alison et al., 2015; Alison, Alison, Noone, Elntib, & Christiansen, 2013; Alison, Alison, Noone, Elntib, Waring, & Christiansen, 2014; Constanzo & Gerrity, 2009; Evans et al., 2013; Evans et al., 2014; Meissner, Redlich, Bhatt, & Brandon, 2012). Psychologically oriented techniques are designed to lower the prisoner's resistance by first establishing rapport and then analyzing the suspect's psyche to determine the best way to gain cooperation (CIA, 2002; Parry & White, 2002).

Conclusions and Implications for Future Research

Various aspects of the "dark side" of persuasive situations such as interrogations and how they relate to deception need further development. More specific to the present review's focus, the deceptive communication within interrogation contexts remains an area not only ready for further research, but possibly an area needing greater scrutiny and understanding in the immediate future.

Research is needed pertaining to interethnic and intercultural interactions between suspects and interrogators. Increased understanding of effective intercultural interrogations would prove invaluable to law enforcement and intelligence agents and their training programs. Further analyses of how interrogations are utilized globally would provide useful frameworks for understanding practices within specific countries. A direct comparison of the interrogation techniques practiced by different countries, with how suspects of different backgrounds respond to being interrogated, would help address these issues. Given the controversial use of "racial profiling" by law

enforcement, further research is also needed to examine how a suspect's race and ethnicity affects the way he or she is interrogated. Of course, verbal and nonverbal behaviors across cultures vary, risking communicative misinterpretations in such situations. Understanding these processes can only increase the fruitfulness of the interrogation process and its outcomes.

A look into the interrogator's perspective is also essential. Specifically, research should further examine the formal preparation interrogators undergo to effectively communicate with suspects (e.g., Cleary & Warner, 2016). Of special concern is the general lack of formal interrogation training received by law enforcement members, and the fact that such training often includes psychologically coercive techniques to be used on both adults and juveniles alike (Cleary & Warner, 2016). Another crucial part of interrogation needing investigation is the degree to which false confessions hinder successful prosecution. Interrogation practices must be modified to reduce the frequency of false confessions, thereby also reducing "interrogation-induced miscarriages of justice" (Cleary & Warner, 2016, p. 270).

At its core, interrogations are meant to extract truthful statements and overcome deceptive attempts. Deception may exist on both sides of the interview table. The tactics used in the interrogation process can be problematic and even inhumane in certain contexts, but while the ethical and humane aspects of these methods may get debated, the evidence shows the methods themselves are actually ineffective, providing reason enough to abandon them.

While police and other officials use methods that might seem cruel or unpleasant to some individuals outside of law enforcement, the use of these tactics is, at times, deemed necessary by law enforcement due to national security and public safety concerns (e.g., Lonky, 2017; Wright, 2011). Yet, interrogation confessions derived under such dire circumstances historically received case-by-case approvals by the courts and provided opportunities for "unconstitutional conduct" by law enforcement (Lonky, 2017, p. 398). Ultimately, this led to legal rulings such as *Miranda v. Arizona* in 1966 to protect citizens from improper interrogations.

Harsher techniques are often believed to be effective by law enforcement members with the least training and experience (Ghosh, 2009). The more highly trained and skilled the interrogator, the less likely they are to use physical and extreme psychological interrogation techniques, primarily because these techniques have been largely ineffective (Alison et al., 2015; Bloche, 2017; Davis, 2014a; Ghosh, 2009). That is, these dark side interrogation techniques not only produce problematic results, but also are largely ineffective at gaining truthful information and detecting deception (Davis, 2014a). As such, law enforcement should recognize the superior effectiveness of psychological interrogation techniques that produce valid information. Continuing to use such ineffective interrogation methods wastes valuable law enforcement, military, and intelligence resources by spending time following

false leads, time in courts with false confessions, and time spent in extended interrogations when more brief and engaging interviews yield superior information.

One of the basic elements of deception detection is the difference between spontaneous lies and those able to be practiced, such as planned and practiced lies (Greene, O'Hair, Cody, & Yen, 1985; O'Hair, Cody, & McLaughlin, 1981). As extended interrogations wear on, suspects' lies become more engrained and less spontaneous, losing valuable interview opportunities. Just as in the case of terrorists receiving training to resist disclosure during interrogation and even torture, those criminals with past experience in police stations, law enforcement interviews, and interrogations, are the *least* likely to be affected by deceptive interrogation techniques, seeing them for the false tactics they are. Conversely, innocent suspects with minimal law enforcement encounters, who view police as legitimate authority and protectors of the innocent, are the most likely to be adversely affected by these same deceptive interrogation techniques, potentially producing false confessions and incarcerating innocent civilians. Law enforcement members should be trained to avoid inducing false confessions (Cassell, 1999), be able to identify what causes them, and know how to recognize them (Leo & Ofshe, 1998). False confessions often further skew the public's negative perceptions of the interrogation process, thereby placing negative connotations on law enforcement agencies' persuasion and deceptive techniques, making legitimate interrogations less tolerated and supported by their social and legal communities.

Several constitutional laws, including the Fifth, Sixth, and portions of the Fourteenth Amendments, protect the rights of US citizens against false confessions, though non-citizens are notably often not extended the same protections by these laws, or to the same extent. The continued use of harsh interrogation techniques suggests to the rest of the world that their use is either effective or somehow justified. Either of these is problematic given their ineffectiveness, but using them also helps other countries to justify reciprocation of their use against US citizens. Collectively, the key to whether harsh interrogation techniques should be tolerated is whether the outcomes are justified in terms of societal cost and what is lost in a free, fair, and open society.

References

Alison, L., Alison, E., Noone, G., Elntib, S., & Christiansen, P. (2013). Why tough tactics fail and rapport gets results: Observing rapport-based interpersonal techniques (ORBIT) to generate useful information from terrorists. *Psychology, Public Policy, & Law, 19,* 411–431. https://doi.org/10.1037/a0034564.

Alison, L., Alison, E., Noone, G., Elntib, S., Waring, S., & Christiansen, P. (2014). The efficacy of rapport-based techniques for minimizing counter-interrogation tactics amongst a field sample of terrorists. *Psychology, Public Policy, & Law, 20,* 421–430. https://doi.org/10.1037/law0000021.

Alison, L., Giles, S., & McGuire, G. (2015, September). Blood from a stone: Why rapport works and torture doesn't in 'enhanced' interrogations. *Investigative Interviewing: Research & Practice*, *7*(2), 5–23.

Alschuler, A. W. (1997). Symposium on coercion: An interdisciplinary examination of coercion, exploitation, and the law: II. Coerced confessions: Constraint and confession. *Denver University Law Review, 74,* 957–978.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274. https://doi.org/10.1037/0033-2909.111.2.256.

Aubry, A. S., & Caputo, R. R. (1980). *Criminal interrogation*. Springfield, IL: Thomas Books.

Balfe, M. (2018). The chiral nature of the enhanced interrogation programme. *Aggression and Violent Behavior, 41,* 32–47. https://doi.org/10.1016/j.avb.2018.05.011.

Blagrove, M. (1996). Effects of length of sleep deprivation on interrogative suggestibility. *Journal of Experimental Psychology: Applied, 2*(1), 48–59. https://doi.org/10.1037/1076-898X.2.1.48.

Bloche, M. G. (2017). Toward a science of torture? *Texas Law Review, 95,* 1329–1355.

Bowden, M. (2003, October). The dark art of interrogation. *The Atlantic Monthly*, pp. 51–70.

Bravin, J., & Fields, G. (2003, March 4). How do U.S. interrogators make a captured terrorist talk? *Wall Street Journal*, p. B.1.

Buller, D. B., & Burgoon, J. K. (1994). Deception: Strategic and nonstrategic communication. In J. A. Daly & J. M. Wiemann (Eds.), *Strategic interpersonal communication* (pp. 191–223). Hillsdale, NJ: Lawrence Erlbaum Associates.

Caplan, G. M. (1985). Questioning Miranda. *Vanderbilt Law Review, 38,* 1417–1476.

Cassell, P. G. (1998, August). Handcuffing the cops: Miranda's harmful effects on law enforcement. *NCPA Policy Report 218*. http://www.ncpathinktank.org/pub/st218.

Cassell, P. G. (1999). The guilty and the "innocent": An examination of alleged cases of wrongful conviction from false confessions. *Harvard Journal of Law & Public Policy, 22,* 526–590. https://doi.org/10.2139/ssrn.161470.

Cassell, P. G., & Hayman, B. S. (1996). Dialogue on Miranda: Police interrogation in the 1990's: An empirical study of the effects of Miranda. *UCLA Law Review, 43,* 839–931.

CIA Interrogations Enter Gray Areas/Terror Suspects Face Sleep Deprivation, Mind Tricks, Brutal Foreign Agents: [3 STAR Edition]. (2002, December 26). *Houston Chronicle*, p. 30.

Cialdini, R. B. (1998). *Influence: The psychology of persuasion*. New York: Quill.

Cleary, H., & Warner, T. C. (2016). Police training in interviewing and interrogation methods: A comparison of techniques used with adults and Juvenile suspects. *Law and Human Behavior, 40,* 270–284.

Cleary, H., & Warner, T. C. (2017). Parents' knowledge and attitudes about youths' interrogation rights. *Psychology, Crime & Law, 23,* 777–793. https://doi.org/10.1080/1068316X.2017.1324030.

Cody, M. J., Canary, D. J., & Smith, S. W. (1994). Compliance-gaining goals: An inductive analysis of actors' goal types, strategies, and successes. In J. A. Daly & J. M. Wiemann (Eds.), *Strategic interpersonal communication* (pp. 33–90). Hillsdale, NJ: Lawrence Erlbaum Associates.

Condit, C. M. (1997). In praise of eloquent diversity: Gender and rhetoric as public persuasion. *Women's Studies in Communication, 20*(2), 91–117. https://doi.org/10.1080/07491409.1997.10162405.

Connery, D. (1977). *Guilty until proven innocent.* New York, NY: G. P. Putnam's Sons.

Constanzo, M. A., & Gerrity, E. (2009). The effects and effectiveness of using torture as an interrogation device: Using research to inform policy debate. *Social Issues & Policy Review, 3,* 179–210.

Davis, D. C. (2014a). Torture. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 893–897). Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781483306902.n344.

Davis, D. C. (2014b). Neurophysiology. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 704–707). Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781483306902.n265.

Davis, D. C. (2014c). Polygraph. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 774–776). Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781483306902.n295.

Davis, W. L., Leslie, P. J., & Davis, A. B. (2014). *Interviewing, interrogation & communication for law enforcement* (3rd ed.). Bloomington, IL: XLIBRIS.

Dawes, J. (2004). Atrocity and interrogation. *Critical Inquiry, 30,* 249–266. https://doi.org/10.1086/421124.

Dowling, J. L. (1979). *Criminal investigation.* New York, NY: Harcourt Brace Jovanovich.

Dripps, D. A. (1999). Miscarriages of justice and the constitution. *Buffalo Criminal Law Review, 2,* 1–22.

Dripps, D. A. (2003). *About guilt and innocence.* Westport, CT: Praeger.

Eastwood, J., Snook, B., Luther, K., & Freedman, S. (2016). Engineering comprehensive youth interrogation rights null. *New Criminal Law Review, 19*(1), 42–62.

Ends, Means and Barbarity—Torture. (2003, January 11). *The Economist*, pp. 366, 22.

Evans, J., Meissner, C., Ross, A., Houston, K., Russano, M., & Horgan, A. (2013). Obtaining guilty knowledge in human intelligence interrogations: Comparing accusatorial and information gathering approaches with a novel experimental paradigm. *Journal of Applied Research in Memory and Cognition, 2*(2), 83–88. https://doi.org/10.1016/j.jarmac.2013.03.002.

Evans, J. R., Houston, K. A., Meissner, C. A., Ross, A. B., LaBianca, J. R., Woestehoff, S. A., et al. (2014). An empirical evaluation of intelligence-gathering interrogation techniques from the United States Army Field Manual. *Applied Cognitive Psychology*. https://doi.org/10.1002/acp.3065.

Finn, P., & Warrick, J. (2009, March 29). Detainee's harsh treatment foiled no plots. Waterboarding, rough interrogation of Abu Zubaida produced false leads, officials say. *The Washington Post*, p. A01.

Freedman, S., Eastwood, J., Snook, B., & Luther, K. (2014). Safeguarding youth interrogation rights: The effect of grade level and reading complexity of youth waiver forms on the comprehension of legal rights. *Applied Cognitive Psychology, 28,* 427–431.

Geddes, L. (2016, June 18). It could be you. *New Scientist*, pp. 36–39.

Ghosh, B. (2009, June 8). How to make terrorists talk. *Time*, pp. 40–43.

Gordon, N. J., & Fleisher, W. L. (2002). *Effective interviewing and interrogation techniques.* San Diego, CA: Academic Press.

Graham, F. P. (1970). *The due process revolution*. New York, NY: Macmillan Company.

Granhag, P. A., & Strömwall, L. A. (2001a). Deception detection based on repeated interrogations. *Legal and Criminological Psychology, 6,* 85–101.

Granhag, P. A., & Strömwall, L. A. (2001b). Deception detection: Interrogators' and observers' decoding of consecutive statements. *The Journal of Psychology, 135,* 603–621. https://doi.org/10.1080/00223980109603723.

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology, 16,* 243–257. https://doi.org/10.1002/acp.784.

Grano, J. D. (1986). 1986 survey of books relating to the law: II crime, punishment, and criminal procedure: Selling the idea to tell the truth: The professional interrogator and modern confessions law. *Michigan Law Review, 84,* 662–690. https://doi.org/10.2307/1288839.

Grano, J. D. (1993). *Confessions, truth and the law*. Ann Arbor: The University of Michigan Press.

Graessner, S., Gurris, N., & Pross, C. (Eds.). (2001). *At the side of torture survivors: Treating a terrible assault on human dignity* (J. M. Riemer, Trans.). Baltimore, MA: The Johns Hopkins University Press.

Greene, J. O., O'Hair, H. D., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. *Human Communication Research, 11,* 335–364.

Gudjonsson, G. H. (1992). *The psychology of interrogations, confessions and testimony*. New York, NY: Wiley.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. London, UK: Wiley.

Gudjonsson, G., Clare, I., Rutter, S., & Pearse, J. (1993). Persons at risk during interviews in police custody: The identification of vulnerabilities. *The Royal Commission on Criminal Justice, 12,* 1–38. https://doi.org/10.1002/9780470713297.ch3.

Hail-Jares, K., Lowrey-Kinberg, B., Dunn, K., & Gould, J. B. (2017). False rape allegations and wrongful convictions. *Family & Intimate Partner Violence Quarterly, 9*(4), 39–45.

Hancock, L. (2003). False confessions: How they happen. *Columbia Journalism Review, 41*(5), 40–42.

Helm, M. (2003, November 23). High court to reconsider interrogation technique; Critics of "two-step" method used by cops say it flouts Miranda rule, violates suspects' rights. *The Times Union*.

Hentoff, N. (2003, January 27). Torture & the CIA. *Editor and Publisher, 136*(4), 34.

Hess, J. E. (2015). *Interviewing and interrogation for law enforcement* (2nd ed.). New York, NY: Routledge.

Horvath, F. S. (1973). Verbal and nonverbal clues to truth and deception during polygraph examinations. *Journal of Police Science & Administration, 1,* 138–152.

Howe, E. G. (2003, June). Dilemmas in the military medical ethics since 9/11. *Kennedy Institute of Ethics Journal, 13*, 175–189.

Inbau, F. E., Reid, J. E., & Buckley, J. P. (1986). *Criminal interrogation and confessions* (3rd ed.). Baltimore, MD: The Williams & Wilkins Company.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2011). *Criminal interrogation and confessions* (4th ed.). Burlington, MA: Jones and Bartlett Publishers.

International Injustice. (2004, February 21). *The Guardian*. Retrieved from https://www.theguardian.com/world/2004/feb/21/terrorism.foreignpolicy.

Jones, T. M., & Sheets, P. (2009). Torture in the eye of the beholder: Social identity, news coverage, and Abu Ghraib. *Political Communication, 26,* 278–295.

Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist, 52,* 221–233. https://doi.org/10.1037/0003-066X.52.3.221.

Leo, R. A., & Ofshe, R. J. (1998). The consequences of false confessions: Deprivations of liberty and miscarriages of justice in the age of psychological interrogation. *Journal of Criminal Law & Criminology, 88,* 429–496.

Libkuman, T. M., Griffith, J., Wines, W. M., Dickel, M. J., & Doty, K. G. (1998). Individual differences in arousal and accessibility to information in memory. *The Journal of General Psychology, 125,* 367–396. https://doi.org/10.1080/00221309809595344.

Liefaard, T., & van den Brink, Y. (2014). Juveniles' right to counsel during police interrogations: An interdisciplinary analysis of a youth-specific approach, with a particular focus on The Netherlands. *Erasmus Law Review, 7,* 206–218.

Lonky, H. (2017). Revisiting the public safety exception to Miranda for suspected terrorists: Dzhokhar Tsarnaev and the bombing of the 2013 Boston Marathon. *Journal of Criminal Law & Criminology, 107,* 393–420.

Lunbeck, E. (1994). *The psychiatric persuasion: Knowledge, gender, and power in modern America*. Princeton, NJ: Princeton University Press.

Macdonald, J., & Michaud, D. (1987). *The confession*. Denver, CO: Apache Press.

Mackey, C. (2004). *The interrogators: Inside the secret war against Al Qaeda*. New York: Little, Brown & Co.

Magid, L. (2001). Deceptive police interrogation practices: How far is too far? *Michigan Law Review, 99,* 1168–1221. https://doi.org/10.2139/ssrn.272659.

Marshall, L. L. (1994). Physical and psychological abuse. In W. R. Cupach & B. H. Spitzberg (Eds.), *The dark side of interpersonal communication* (pp. 281–311). Hillsdale, NJ: Lawrence Erlbaum Associates.

Massimino, E. (2004). Leading by example? U.S. interrogation of prisoners in the war on terror. *Criminal Justice Ethics, 23*(1), 2, 74–76. https://doi.org/10.1080/0731129X.2004.9992155.

Maver, D. (1996). *Defence strategies and techniques of interrogation: Results of empirical research*. Policing in Central and Eastern Europe: Comparing Firsthand Knowledge with Experience from the West. Slovenia: College of Police and Security Studies.

Meissner, C. A., Redlich, A. D., Bhatt, S., & Brandon, S. E. (2012). Interview and interrogation methods and their effects on true and false confessions. *Campbell Systematic Reviews, 13,* 1–53. http://doi.org/10.4073/csr.2012.13.

Middleton, K. R., Lee, W. E., & Chamberlin, B. F. (2004). *The law of public communication* (6th ed.). New York, NY: Pearson Education.

Nagel, S. S. (Ed.). (1972). *The rights of the accused*. Beverly Hills, CA: Sage.

O'Hair, H. D., & Cody, M. J. (1994). Deception. In W. R. Cupach & B. H. Spitzberg (Eds.), *The dark side of interpersonal communication* (pp. 181–213). Hillsdale, NJ: Lawrence Erlbaum Associates.

O'Hair, H. D., Cody, M. J., & McLaughlin, M. L. (1981). Prepared lies, spontaneous lies, Machiavellianism, and nonverbal communication. *Human Communication Research, 7,* 325–339. http://dx.doi.org/10.1111/j.1468-2958.1981.tb00579.x.

O'Hara, C. E., & O'Hara, G. L. (1998). *Fundamentals of criminal investigation* (5th ed.). Springfield, IL: Charles C. Thomas.

Paine, C. (2004). Bush's "dirty war". *Bulletin of the Atomic Scientists, 60*(6), 78–80.

Panetta, L., & Serraglio, R. (2000, November/December). Training assassins. *NACLA Report on the Americas, 34*(3), 28–29. https://doi.org/10.1080/10714 839.2000.11722618.

Papenfuss, M. (2018, March 15). Edward Snowden takes on Liz Cheney over torture links to Trump's pick for CIA. *Huffington Post.* https://www.huff-ingtonpost.com/entry/snowden-attacks-cheney-over-haspel-torture_us_5aab-197de4b0c33361af21a2.

Parry, J., & White, W. (2002). Interrogating a suspected terrorist. *University of Pittsburgh School of Law, 63,* 743–766.

'Rectal Hydration': Inside the CIA's interrogation of Khalid Sheikh Mohammed (2014, December 9, 6:55 PM ET). *NBC News.*

Riechmann, D. (2018, April 20). Probe clears Haspel on destruction of waterboarding tapes. *Associated Press Wire.*

Ries, E. (2017, November). PTs helping to heal survivors of torture. *PT in Motion, 9*(10), 16–25.

Rejali, D. (2007, December 16). 5 myths about torture and truth. *The Washington Post.*

Sasaki, D. W. (1988). Guarding the guardians: Police trickery and confessions. *Stanford Law Review, 40,* 1593–1616. https://doi.org/10.2307/1228783.

Seiter, J. S., & Gass, R. H. (2004). *Perspectives on persuasion, social influence, and compliance gaining.* New York, NY: Pearson Education.

Shakespeare, W. (1995). *The merchant of Venice.* Mineola, NY: Dover.

Shepard, E. (1991). Ethical interviewing. *Criminological and Legal Psychology, 18,* 46–56.

Shuy, R. W. (1998). *The language of confession, interrogation, and deception.* Thousand Oaks, CA: Sage.

Skolnick, J. H., & Fyfe, J. J. (1993). *Above the law.* New York, NY: Macmillan.

Slobogin, C. (1997). Deceit, pretext, and trickery: Investigate lies by the police. *Oregon Law Review, 76,* 775–816.

Smith, P., Shane, S., Mazzetti, M., & Baker, P. (2015, February 2). Is torture ever justified? A new report on the C.I.A.'s brutal interrogation of terrorism suspects since 9/11 has reignited the debate over the use of torture. *New York Times.* Retrieved from https://www.thefreelibrary.com/Is+torture+ever+justified%3f+A+new+report+on+the+C.I.A.%27s+brutal…-a0399571199.

Solomon, J. (2003, December 1). Saudis use new interrogation technique. *Associated Press.*

Spitzberg, B. H., & Cupach, W. R. (1994). Dark side dénouement. In W. R. Cupach & B. H. Spitzberg (Eds.), *The dark side of interpersonal communication* (pp. 315–320). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spitzberg, B. H., Marshall, L., & Cupach, W. R. (2001). Obsessive relational intrusion, coping, and sexual coercion victimization. *Communication Reports, 14*(1), 19–31. https://doi.org/10.1080/08934210109367733.

Stableford, D. (2018, March 13). Trump CIA pick Gina Haspel helped lead U.S. torture program. *Yahoo News.* https://www.yahoo.com/news/trump-cia-pick-gina-haspel-helped-lead-u-s-torture-program-164726899.html.

Stein, J. (2014, December 15). A tortured past. *Newsweek*, pp. 12–15.

Stewart, J. M., Woody, W. D., & Pulos, S. (2018). The prevalence of false confessions in experimental laboratory simulations: A meta-analysis. *Behavioral Sciences and the Law, 36*, 12–31. http://dx.doi.org/10.1002/bsl.2327.

Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors, and judges. *Psychology, Crime, & Law, 9*, 19–36. https://doi.org/10.1080/10683160308138.

Trenholm, S. (2005). *Thinking through communication* (4th ed.). Boston, MA: Pearson Education.

Wakefield, H., & Underwager, R. (1998). Coerced or nonvoluntary confessions. *Behavioral Sciences & the Law, 16*, 423–440.

Walkley, J. (1987). *Police interrogation: A handbook for investigators.* Cambridge, UK: Black Bear Press Limited.

White, W. S. (1997). False confessions and the constitution: Safeguards against untrustworthy confessions. *Harvard Civil Rights—Civil Liberties Law Review, 32*, 105–149.

White, W. S. (1998). What is an involuntary confession? *Rutgers Law Review, 50*, 2001–2057.

White, W. S. (2001a). Miranda's failure to restrain pernicious interrogation practices. *Michigan Law Review, 99*, 1211–1248. https://doi.org/10.2307/1290530.

White, W. S. (2001b). *Miranda's waning protections.* Ann Arbor: The University of Michigan Press.

Woody, W. D., & Forrest, K. D. (2009). Effects of false-evidence ploys and expert testimony on jurors' verdicts, recommended sentences, and perceptions of confession evidence. *Behavioral Sciences & the Law, 27*, 333–360.

Woody, W. D., Forrest, K. D., & Stewart, J. M. (2011). False confessions: The role of police deception in interrogation and jurors' perceptions of the techniques and their outcomes. In A. E. Hasselm (Ed.), *Crime: Causes, types and victims* (pp. 1–37). New York: Nova Publishing.

Wright, J. (2011). Mirandizing terrorists? An empirical analysis of the public safety exception. *Columbia Law Review, 111*, 1296–1331.

Yeschke, C. L. (2003). *The art of investigative interviewing* (2nd ed.). New York, NY: Elsevier Science.

York, B. (2004). A tortured debate. *National Review, 56*(13), 33–36.

Young, D. (1996). Unnecessary evil: Police lying in interrogations. *Connecticut Law Review, 28*, 425–469.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking.* New York: Cambridge University Press.

Zulawski, D. E., & Wicklander, D. E. (1992). *Practical aspects of interview and interrogation.* New York, NY: Elsevier Science.

Contexts of Deceptive Communication: Mass Mediated Communication

Deception and the Social Good in Mass Communication

Seow Ting Lee

Persuasion, albeit a controversial subject, is one of the key processes and outcomes in mass communication. To persuade, many mass media messages consciously or otherwise seek to shape, alter, and reinforce the perception, cognition, and behavior of large groups of people and multiple publics. One of the fundamental tools of persuasion is deception, or a message knowingly transmitted by a sender to foster a false belief in the receiver (Buller & Burgoon, 1996; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996). At the heart of deception lies the notion of intent. A message is considered deceptive if it is motivated by the sender's intent to mislead the receiver, thus ruling out honest error.

This chapter reviews the literature on deception in mass communication, specifically in the fields of public relations (PR), journalism, advertising, and health communication. The four fields, with a shared grounding and yet distinct goals and functions, offer a rich contextual locus for explicating deception as a theoretical and professional construct. In three cognate fields—public relations, advertising, and health communication, communicators by virtue of their job functions and goals assume the stance of professional advocates for organizations, groups, and individuals by using selective truth to persuade their audiences and publics. Truth is central to journalism's professional self-conception but is shaped by distinct philosophical antecedents of public service and the public good that demand the use of deception in news gathering.

S. T. Lee (✉)
Department of Advertising, Public Relations and Media Design,
University of Colorado Boulder, Boulder, CO, USA
e-mail: seowting.lee@Colorado.EDU

To communicate—and to persuade successfully, one has to be believed. Deception is strategic. The opposite of deception is truth-telling, the closest to a universal value. If so, is deception always wrong? Are there situations wherein deception may be justified? Selective application of the truth creates a false version of reality but not all deception is wrong, and deception may be beneficial in some contexts (Bok, 1978; Bonhoeffer, 1965). Italian philosopher and theologian Thomas Aquinas opined that all lies are wrong but maintained that not all lies are sins. The discussion of the rightness or the wrongness of deception inherently centers on tensions between two prevailing schools of ethical thought. From a deontological perspective, truth-telling is a duty. Truthfulness is viewed as a virtue expressed as part of fundamental obligations, including respect for autonomy, justice, beneficence, and non-maleficence. To not tell the truth is to fail to respect others and their autonomy, to violate implicit contracts, and to damage relationships based on trust. From a deontological perspective, the moral rightness of an action is assessed based on the act rather than its consequences.

One of the strongest deontological treatises against deception was offered by German philosopher Immanuel Kant. In *Grounding for the Metaphysics of Morals: On a Supposed Right to Lie Because of Philanthropic Concerns*, Kant (1785/1993), who defined a lie as any intentional statement that is untrue, viewed lying as an affront to human dignity and mankind. By contrast, the teleological school of thought emphasizing outcomes treats deception as simply another purposeful, goal-directed endeavor. By focusing on the benefits of deception, one could justify the use of deception to maximize the greatest good for the greatest number within a society's limited resources. In Greek philosopher Aristotle's scheme of virtue ethics and vision of social good, truth-telling is what a person of good character would do, although the same person may lie to be compassionate and courageous.

Deception may help mass communicators achieve professional goals, but it has many far-reaching repercussions on public opinion and trust, as well as unintended consequences that compromise message efficacy and relational successes with audiences and publics. Bok (1978) advanced a strong moral presumption against deception in her seminal book *Lying: Moral Choice in Public and Private Life*. Bok's *principle of veracity* asks if you would like to live in a world in which truth-telling is not a norm. In such a world, you could never trust anything you were told or anything that you read. Such a world would require you to find out everything for yourself first-hand—and render redundant many of the professional functions performed by journalists, public relations practitioners, advertisers, and the health communicators. Bok defined a lie as a statement, believed by the liar to be false, made to another person with the intention to deceive him or her. Bok's *principle of veracity* suggests that a lie is advantageous only in situations where people will believe it. In other words, lies can only work in societies where truth-telling prevails. The principle involves two steps: First, you must believe that

you personally benefit from a system that you want others to do their part in maintaining. Second, reciprocity or fair play must exist, requiring you to do your part in maintaining the system if others are doing their part.

It is easy to disparage mass communicators for their use of deceptive practices, but deception is systemic in human interactions. Deception occurs in at least one-quarter of all conversations (Buller & Burgoon, 1996; DePaulo et al., 1996; Turner, Edgley, & Olmstead, 1975). The fact that people deceive however does not make deception acceptable, but the rampant use of deception serves to highlight its complex nature. Mieth (1997) described truth-telling as a "basic norm" but acknowledged that humans are quick to "invoke at one moment the norm of truthfulness and at the next moment the right to lie, depending on circumstances and context" (p. 87). However, the fact that one is expected to justify any departure from truth-telling implies that deception is a prima facie wrong.

The discussion of deception also necessarily entails a recognition of the broad spectrum nature of truth. For millennia, examinations of truth have dominated philosophy, with one of the earliest examples being Plato's *Allegory of the Cave* that discussed the role of human perception in the construction of knowledge and truth. Patterson and Wilkins (2013) showed how truth, or what was considered to correspond to facts and reality, changed over time, beginning from truth as memory and oral discourse to the truth as viewed in the Platonic, Medieval, Miltonic, Enlightenment, and Pragmatic traditions. Although truth is typically used in contemporary society to refer to a state of correspondence with reality or fact, philosophers and theologians continue to discuss and debate this illusive concept through divergent views of truth including relativism, which uses different types of relativity to show that absolute truth is hard to establish.

One can deceive not only through intentional fabrication of facts but also through the intentional withholding of facts. Many philosophers (e.g., Chisholm & Feehan, 1977; Fried, 1978) considered deception by omission to be less egregious than deception by commission or the act of actively misleading a person with an overt statement that is untrue (i.e., a lie). Elliot and Culver (1992) however contended that some forms of deception by omission could be more problematic than deception by commission.

Given the importance of deception as a theoretical construct in mass communication, it is surprising that deception has not attracted commensurate attention from mass communication scholars. A review of the literature reveals a sparse and fragmented body of literature. Although we cannot ignore the marked differences in philosophical antecedents, professional goals and functions between and among public relations, journalism, advertising, and health communication, these four fields broadly share a common reliance on truth-telling as a key foundation of effective mass communication. Deaver (1990) is one of the first mass communication scholars to define and to operationalize truth. Deaver observed that "communication of all sorts is

passed off as 'truth,' when in fact it is a collection of truth, half-truth and untruth" and "[c]ommunicators often deal in varying degrees of truth" (p. 168). Instead of providing a definition of truth, Deaver operationalized truth through application in various mass communication contexts based on four categories of truth and untruths in mass communication according to:

1. intent to inform, accurately and fully with no apparent bias (e.g., facts and information, journalistic news).
2. intent to persuade by using selective information, truth but not the whole truth (e.g., public relations writing, editorials/columns, advertising copy).
3. nontruths told without intent to deceive (e.g., fiction, honest error).
4. intent to deceive, even if for purposes thought justifiable (e.g., deceit, white lies, blatant lies).

Persuasion, according to Deaver, does not permit the telling of untruths but merely introduces the element of selectivity in communication, with the assumption that truth is still achievable. Selective information, although not the whole truth, can be used to ethically persuade the audience without compromising truth-telling. The construction of truth in public relations, advertising, and health communication thus falls into the category of persuasive messages that are ethical. Deaver drew the line at blatant lies or untruths "communicated with no redeeming purpose, with only an unjustifiable intent to deceive" (p. 176). Such lies might be justified only by the teleological argument that the lie serves to avoid a tragic outcome. Although Deaver's model provided a normative framework for assessing the ethics of the communicator within each of the degrees of truthfulness, his model remained largely untested empirically.

Public Relations: A Troubled Relationship with Truth

Among the four mass communication fields, criticisms of deception tend to focus on public relations' troubled relationship with truth. In the eyes of the public, the practice of public relations is popularly associated with propaganda and spin—hence, the terms "spin" and "spin doctors." Due to its historical associations with manipulation and press agentry, public relations continues to struggle with poor reputational standing especially when judged using the truth-telling yardsticks of journalism. The pejorative characterization and dismissal of public relations as propaganda "simplistically defines propaganda as lying which, by implication, places it in opposition to truth" (Weaver, Motion, & Roper, 2006, p. 2). Propaganda as a construct for explaining deception in public relations can be attributed to a continued reliance on the transmission model of communication as a theorizing framework, viewing the receiver as a vulnerable and persuadable lot. However, the use of selective truths through persuasion's liberal theoretical roots aligns with ideals of freedom of media and of expression. Public relations practitioners—akin to lawyers—are tasked with advancing their clients' interest in an adversarial system

grounded in debate and exchange of ideas that underpin democratic processes in most contemporary societies. In the US, messages of selective truth used by public relations practitioners are protected by law, grounded in the marketplace of ideas' mechanism of self-righting truth where "[t]ruth should prevail in a market-like struggle where superior ideas vanquish the inferiors and achieve audience acceptance" (Lentz, 1996, p. 1). Reflecting the client- and organization-driven work of public relations, and its roles in managing relationships between organizations and its publics, and its advocacy roles, much of the literature on deception is focused on explicating the tensions between public relations practitioners' obligations to their clients and their organizations versus their obligations to their publics and society at large.

Paradoxically, there is a consensus in the scholarly literature that deception is unacceptable (Baker, 2008; Bivins, 2006; Bowen, 2006, 2008; Fitzpatrick & Gauthier, 2001; Lee, 2011a; Lee & Cheng, 2012; L'Etang, 1994; Parsons, 2016; Patterson & Wilkins, 2013). Parsons (2016) called truth "a cornerstone of public relations as a communication industry" (p. 18). Further, Parsons stated:

> So it seems that telling the truth, although often deemed to be a casualty in the search for new and better ways to disseminate messages and persuade publics, is an important aspect of ethical public relations. Defining what the truth is in public relations, just as in other aspects of our lives, however, is a challenge. (p. 15)

Parsons offered this guideline: "If telling the truth outright is likely to harm one or more publics, then it is reasonable to conclude that it is probably more ethical to avoid full disclosure" (p. 15).

Seib and Fitzpatrick (1995) noted that certainty about truth is often elusive and is influenced by opinion, completeness of information, interpretation, and perception. What is truth? This is one of the oldest questions. It is perpetually difficult to answer but a number of ethical issues in public relations relate to truth directly and indirectly, for example in issues about accuracy, withholding of information, client confidentiality, conflict of interest, media relations, transparency on the Internet, and avoidance of harm.

Lee and Cheng (2011) in a survey of 350 public relations practitioners in the US found that the most pressing ethical challenge is telling the truth, followed by maintaining factual accuracy of information released to the public, and commitment to public interest. Lee (2011a), using in-depth interviews with public relations executives, found that truthfulness is a key characteristic of an ethical leader. Interviewees typically linked truthfulness with credibility and considered truth-telling to be "one of the basic values that underlie everything" (p. 59). Truth-telling is tempered by confidentiality in not revealing clients' trade secrets, with a caveat reflecting the primacy of public interest or only if confidentiality does not inflict harm on the public.

In public relations, ethics is heavily dependent on codes of ethics that outline behaviors that are discouraged and encouraged. Codes represent a

guide to professional conduct and a framework for understanding the moral obligations of practitioners. In the Public Relations Society of America's (PRSA) Code of Ethics, which gives prominence to truth, members are urged to avoid deceptive practices. The PRSA is the largest association of public relations and communications professionals in the US. According to the code, "We adhere to the highest standards of accuracy and truth in advancing the interests of those we represent and in communicating with the public." Under the Free Flow of Information provision, the core principle states that "protecting and advancing the free flow of accurate and truthful information is essential to serving the public interest and contributing to informed decision making in a democratic society" (PRSA code of ethics, n.d.). Guidelines under this principle include "being honest and accurate in all communications" and "acting promptly to correct erroneous communications for which the practitioner is responsible." In the International Association of Business Communicators (IABC) code of ethics, provisions for truth-telling addressed honesty, accuracy, and confidentiality (Code of Ethics for Professional Communicators, 2016):

- I am honest—my actions bring respect for and trust in the communication profession.
- I communicate accurate information and promptly correct any errors.
- I protect confidential information while acting within the law.

Similar exhortations are found in the codes of public relations agencies. One example, from Edelman's Code of Conduct, states the following: "We adhere to the highest standards of accuracy and truth in advancing the interests of those we represent and in communicating with the public"; "We do not intentionally disseminate false or misleading information or omit critical information that is essential to avoid misinformation"; "While a lack of transparency can sometimes yield short-term results, it can call into question the integrity of communications and create an unacceptable level of risk for our reputation and that of our clients"; and "We do not act in a way that may appear to be an attempt to deceive public opinion" (Edelman Code of Ethics and Business Conduct, 2017).

While laudable, codes of ethics are difficult to enforce—and for a complex notion such as truth, provisions in codes are too simplistic or too general for practical everyday use, as demonstrated by the body of literature critical of codes of ethics (e.g., Bivins, 2006; Curtin & Boynton, 2001; Fitzpatrick, 2002). Some scholars (e.g., Curtin & Boynton, 2001; Lee & Cheng, 2010) do not dismiss the usefulness of codes entirely but suggest that codes are more useful for neophytes, in helping them understand their professional moral obligations. Ikonen, Luoma-aho, and Bowen (2017) analyzed 40 codes of ethics in public relations, advertising, and journalism in the US and Finland. Their study focused on the rise of sponsored content globally that blurs the lines between strategic communication, advertising, and journalism.

Their analysis found excessive variation within the codes about transparency of sponsored content.

As public relations theory and research move from a one-way, informational model toward a two-way, relational approach, the notion of dialogue is evolving into an important construct for measuring public relations success and for explicating a clearer framework for understanding and applying deception in public relations. This movement is not inconsistent with the idealistic positioning of the role of public relations in the marketplace of ideas. Dialogue has long been considered by philosophers and rhetoricians to be a primary means for separating truth from falsehood. The earliest explication of dialogue in public relations can be traced to Pearson (1989) who viewed dialogue as a foundation for public relations ethics. Kent and Taylor (2002), who clarified the concept of dialogue in public relations, suggest that dialogue is "honest and forthright," and involves revealing one's position "in spite of the possible value that deception or nondisclosure might have" (p. 29). As public relations evolves and solidifies its relational approach, dialogue could provide a promising path for a revitalized framework of truth-telling.

Journalism: Paradoxes and Double Standards

Journalism deals with gathering and presenting news to provide citizens with information to make the best possible decisions about their lives, communities, and governments. As a field, journalism stands apart professionally in its goals and functions in a democratic society. Journalists' obligation to tell the truth features prominently in journalism codes of ethics, popular culture, and scholarly literature. Journalism's claim to truth "legitimizes journalism's special position as Fourth Estate" (Broersma, 2010, p. 25). According to Broersma, "Though the impossibility of a mimetic and purely objective representation of reality is commonly accepted, it is striking that journalism's claim to truth and authenticity is still so vivid in journalism and in public discourse" (p. 21). Lately, fake news, or the use of fabrication and deliberate misinformation, has generated considerable controversy, especially in the context of American election processes. Allegations of fake news continue to fuel concern and argument over the credibility of American news media and the effects on public trust, reaffirming the primacy of truth-telling as a journalistic and moral obligation.

The discourse on ethics in mass communication often juxtaposes PR and journalism. In contrast to the pejorative labeling of public relations practitioners as spin doctors and propagandists, theoretical and professional conceptions of journalism have been dominated by what Fisher (2016) called "the idealized role of the 'watchdog' reporter as defender of democracy and seeker of truth" (p. 665). Public relations will always be unethical when measured against journalism. Barney and Black (1994) suggested that a "moral confusion" arose when "[p]ublic relations practitioners are torn between two distinct heritages: the objectivity ethic of the journalist, and the persuasion ethic

of the advocate" (p. 233). Fisher (2016), in a comparative study of public relations and journalistic roles, offered a typology of ten approaches to truth-telling adopted by journalists in their new role as political media advisers. The findings, based on investigating how journalists transition between two roles traditionally portrayed as binary opposites in their approaches to truth-telling, demonstrated that truth is malleable in both communications roles.

Patterson and Wilkins noted that truth-telling is fundamental to journalism, in that journalists "have a greater responsibility to tell the truth than most professions" (p. 74). Singer (2007) argued that commitments to truth and to "transparency," or public accountability, are two central normative aspects of professional journalism. However, journalism is not free of persuasion. Although the persuasive intent of journalists is not as explicit as that of their colleagues in advertising, public relations, and health communication, journalists can rely on a powerful arsenal for disseminating selective truths—intentionally or otherwise through the framing of facts and figures, and the use of emotive language in straight news. More conspicuously lacking in objectivity are news editorials and op-eds that are meant to persuade.

Deception in journalism has been studied extensively in mass communication. Elliot and Culver (1992) defined journalistic deception as an act of communicating messages verbally (i.e., a lie) or nonverbally through the withholding of information with the intention to initiate or sustain a false belief. The literature implicitly separates journalistic deception into two forms: deception in newsgathering and deception in news reporting. While journalists would take an absolutist view in rejecting deception in news reporting, they are more likely to adopt a non-absolutist approach to deception in newsgathering because some deception is indispensable in newsgathering. Kieran (1997) remarked, "Paradoxically, we demand that journalists tell the truth and yet, to get at the truth, they may have to lie" (p. 66). Lee (2004) found that journalists viewed deception as a strategy for gathering information that is of vital public service, including exposing wrongdoing by government officials or informing consumers of fraudulent practices—used ideally as a last resort, when all other means have been exhausted. A utilitarian calculus is used as journalists weigh the harm done to a few individuals—by deceiving newsmakers or news sources—against the larger benefit accrued to society. Some forms of deception, such as a reporter not identifying himself by going undercover, can in fact enhance the objectivity of a news story.

The absolute rejection of deception in news reporting is consistent with the profession's devotion to accuracy and truthfulness as a professional standard. Bagdikian (2000) argued that devotion to accuracy is one of the main strengths of American journalism, and possibly the only professional standard that has not been eroded by the business's economic pressures, with conventions against fictionalizing and factual inaccuracy being "strong and widespread" (p. 212). However, when journalists resort to deception in news reporting, they are more likely to apply the Kantian perspective, in what

would appear to the unaware publics as mixed signals or a clear display of hypocrisy.

Lee (2004) suggested that journalistic deception is an occupational construct shaped by professional demands. Like lies by the undercover police, the journalists' use and assessment of deception are a function of a "negotiated occupational order" (Hunt & Manning, 1991). According to Lee (2004), journalistic deception is shaped by three tacit rules:

1. *News audience vs. Newsmakers*: It is more acceptable to deceive newsmakers than news audiences. The latter is seen as a means to an end, the end being the news audiences that journalists believe they ultimately serve.
2. *The perceived character of the deceived*: It is more acceptable to deceive a bad person than a good person. Deception is viewed not only as a means to an end but also a retaliation for injustice, a motive reinforced by the altruistic values that lead journalists to view themselves as champions of truth and social justice.
3. *Omission vs. Commission*: Deception by omission is considered more acceptable than deception by commission. Acts that involve active falsification of information such as fabrication, quote tampering, photograph manipulation, lying, and impersonation are differentiated from acts involving omission of information—not identifying oneself as a journalist, withholding a story, and using hidden cameras. There is a higher tolerance of deception by omission because it is easier to conceal than to falsify, and a lower risk of being caught because a fictional account does not have to be invented in advance. Concealment also is passive and generates less guilt. Deception by omission also is easier to explain because the deceiver could offer excuses ranging from ignorance, memory lapse, or intention to share information at a more opportune time.

With its distinct set of tacit rules and subtleties, journalistic deception as an occupational construct is shaped by a complex interplay of values and norms central to the broader moral framework in which the profession is located. As the tacit rules are derived from a negotiated meaning within the profession, they are easily lost on outsiders. Given the double standards, mixed signals, and differential treatment accorded to newsmakers versus news audiences, it is unsurprising that when journalistic deception is subjected to public scrutiny, internal rules fall short in the eyes of the public and quickly fan public outrage. In today's complex media landscape, journalism's commitments to truth are increasingly being challenged. More so than ever, the definitions of journalism's professional constructs are being reinterpreted and any departure from truth-telling is becoming more visible—and more vulnerable to public vilification.

ADVERTISING: SELF-REGULATED EXPRESSIONS OF TRUTH

As advertisers seek to persuade consumers to purchase a specific product or service, critiques of advertising are intrinsically fixated on truth. In most societies, truth in advertising standards mandates that advertisers must have evidence to support the express and implied claims made about products and services featured in advertisements. In the US, under the US Federal Trade Commission (FTC) Act, advertising must be truthful, non-deceptive and its claims must be supported by proof. The FTC was established in 1914 to regulate unfair advertising but soon expanded its jurisdiction over misleading advertising that deceived consumers about the attributes, characteristics, and performance of products and services. Petty (2015), who examined the history of US advertising regulation, found that early US courts were reluctant to find advertising dishonest and showed how the advertising industry began embracing regulation to enhance advertising credibility in the 1800s.

In the UK, advertising codes require that advertisers hold evidence to prove the claims that they make before they are published or aired. However, in the context of the millions of advertising messages disseminated, not many cases of deception actually escalate to the level of state interference because the prevailing approach is still self-regulation by industry. The Advertising Standards Authority (ASA) is the country's independent regulator of advertising across all media (ASA Advertising Codes, n.d.). ASA members come from the advertising industry including advertisers, media owners, and advertising agencies. The country's advertising messages are regulated through a system of self-regulation based on funding from the advertising industry, and co-regulation. According to ASA, two forms of regulation are used:

- Self-regulation: The ad industry writes the rules (through CAP) that advertisers have to stick to. Self-regulation covers non-broadcast advertising, including newspapers, posters, websites, social media, cinema, emails, leaflets, and billboards.
- Co-regulation: ASA has a contractual arrangement with the UK communications regulator, Ofcom that gives ASA responsibility on a day-to-day basis to regulate TV and radio advertising. In 2014, Ofcom renewed its contract with ASA for another 10 years.

In 2016, ASA resolved over 29,000 complaints relating to just under 16,000 ads. In addition, it resolved 5425 cases on its own initiative. As a result, 4584 ads were either changed or removed. About 97% of the complaints that ASA received in 2016 were from members of the public, with 72% of the complaints focused on potentially misleading ads.

There is a well-established and substantial body of research on deception in marketing and psychology focused on a teleological perspective (e.g., Armstrong, Gurol, & Russ, 1979; Darke & Ritchie, 2007; Gardner, 1975; Hyman, Tansey, & Clark, 1994; Ullah & Hussain, 2015; Wilkins,

Beckenuyte, & Butt, 2016; Xie, 2016; Xie, Madrigal, & Boush, 2015). The use of deception also is one of the main topics in advertising research and is mostly focused on three main threads: the extent to which consumers would be deceived by claims, consumer reactions to deception, and how deceptive ads affect the efficacy of ads and consequently their impact on customer purchasing intentions. Darke and Ritchie (2007) showed that deceptive advertising negatively affects people's responses to subsequent advertising from the same source and undermines the credibility of advertising in general. Boush, Friestad, and Wright (2015) addressed the psychology of deceptive persuasion in the marketplace and the psychology of consumer self-protection. Xie et al. (2015) studied the effect of anticipated harm on consumer brand attitudes and purchase intentions from that of perceived deception. They found that greater perceived harm increased the ease of detection of deception in ads that partially explains consumers' negative reactions to deceptive advertising.

Although there is consensus that deception is damaging to advertising, few studies however have explicated communicator responsibility in advertising. Baker and Martinson's (2001) TARES is the first theoretical framework to explicate the notion of communication practitioner accountability toward the message receiver in persuasive communication. The TARES is a general framework for persuasive messages including advertising and public relations. Through a five-part test, TARES establishes ethical boundaries for persuasive messages. The five interconnected principles that form the acronym TARES are *Truthfulness* of the message, *Authenticity* of the persuader, *Respect* for the person being persuaded, *Equity* of the persuasive appeal, and *Social Responsibility* for the common good. According to Baker and Martinson, "Although professional persuasion is a means to an instrumental end, ethical persuasion must rest on or serve a deeper, morally based final (or relative) end" (p. 172). As a whole, these five principles "comprise the legitimate end of professional persuasive communications and that these communications are ethical and morally justified if they adhere to the principles of truthfulness, authenticity, respect, equity, and social responsibility" (p. 172). As a normative framework, TARES is rooted in a deontological approach over the teleological approach by suggesting that the persuasive message must be evaluated in itself and not merely based on consequences. According to Baker and Martinson (2001), "advertisers would play an increasingly dysfunctional role in the communications process if means continue to be confused with ends in professional persuasive communications" (p. 148).

The first study to apply the TARES framework and to operationalize its five principles for persuasive communication was by Lee and Cheng (2010), who studied the ethicality of US antismoking ads. Through a content analysis of 826 US TV ads, Lee and Cheng, who developed a coding protocol of 19 items to operationalize the five TARES ethical principles, found that the TARES held up well in antismoking ads, with one-third of ads fulfilling the 19 items. Specifically, they used eight items to assess the visual and

verbal content of ads for elements of truth-telling, omission, exaggeration, and intention to mislead or to deceive. The operationalization of *Truthfulness* is multifaceted, as it goes beyond the veracity of the information presented to also assess omission of information, which can also be deceptive. Ads are inherently time- or space-limited thus restricting the amount of information that could be disseminated. Many ads communicate only part of the truth but not all omissions are considered to be deceptive. For deception to occur, there must exist the intent to deceive. Many ads also contain exaggerations or fluff, but an exaggeration is not deceptive unless there is intent to mislead. Patterson and Wilkins (2002) presented the example of a Cheerios commercial that omitted the fact that there are other components of a heart-healthy lifestyle, and that other breakfast cereals are equally healthful, but the commercial does not lead the consumer to make false assumptions and bad choices. Using the TARES framework, the Cheerios commercial would meet the *Truthfulness* principle although it communicated only part of the truth.

In a content analysis examining the ethicality of fast-food advertising including truthfulness in 380 Burger King and McDonald's ads, Lee and Nguyen (2013) found that among the five TARES principles, *Truthfulness* is weakest, with only 37 ads (9.74%) fulfilling all eight *Truthfulnes*s items. Ads targeting children and teenagers are associated with lower truthfulness than ads targeting adults and the general audience, lending empirical support to the large body of literature critical of fast-food advertising's exploitative approach of targeting the young.

Health Communication: Noble Causes Obscure Issues with Truth

Health communicators including those who work in health promotion and social marketing to create and use products, programs, or interventions to promote health changes in individuals and communities have jobs that put them firmly in the category of professional persuaders like advertisers and public relations practitioners. Like advertising messages, health communication messages are not merely informational but also are persuasive as a form of strategic, goal-driven communication seeking to engender positive changes in people's lives by promoting healthful attitudes and behaviors for a larger public good. In doing so, public health campaigns seek to persuade audiences to adopt particular beliefs or pursue specific courses of action, for example exercising more, quitting smoking, or consuming more vegetables and fruits. Attempts to persuade in health communication can take various forms, ranging from simple rhetorical maneuvering through wordsmithing that can serve to mislead, to coercive health messages that appeal to fears or prejudices by exaggerating certain claims—similar to what advertisements do to influence consumers to purchase a product or service. The difference however is between how audiences perceive commercial advertising and health messages

disseminated by health authorities who are typically well positioned as credible sources of health information.

As in the other mass communication fields, beneficence and avoidance of harm are two prominent guiding principles in health communication. In message content, health communicators should provide their audience with truthful information. In addition, health communicators should respect the audience's autonomy by refraining from deceptive, misleading, manipulative, or coercive techniques. These two interconnected principles underpin an approach to persuasion that, because of its emphasis on protecting autonomy, was described by Hove (2014) as the liberal model of ethical persuasion.

In reality, the discourse in health communication ethics has been dominated by a teleological perspective, relying on a utilitarian reasoning of valuing outcomes and maximizing benefit over harm. A good example is the use of inflated fear appeals in public health messages that seek to frighten the audience or elicit trepidation of consequences of not conforming to behaviors espoused by health authorities. The reliance on teleological ethics, by focusing on consequences as the main determinant of a message's ethicality, has been questioned by some scholars who call for public health messages to be evaluated for its intrinsic moral worth rather than outcome alone (e.g., Guttman, 1997, 2000, 2003; Kirby & Andreasen, 2001; Kirklin, 2007a, 2007b; Lee, 2011b). Although there may be circumstances wherein the larger public good should prevail over the needs and rights of individuals, for instance required vaccinations, pandemics, or quarantines for infectious diseases, it is difficult to justify public health messages that are carried out deceptively from a purely teleological reasoning. Hove (2014), however, argued that respect for autonomy of audiences by persuading them using truthful substantive information is a flawed approach, because "to account for circumstances when respecting autonomy might take a back seat to other ethical considerations, a comprehensive framework for the ethics of health communication needs to acknowledge types of communication that aim to do something other than provide substantive information (disclosive and directive communication), and subjective circumstances when people are not motivated to process information or make active choices (low processing motivation and decision aversion)" (p. 134). Such reasoning is consistent with the strategic and utilitarian approach to the practice of health communication and health communication scholarship focused on problem-solving, by identifying, examining, and overcoming obstacles to better health. For example, the literature focuses on how to enhance communication's impact in health promotion through tangible, measurable outcomes such as drops in smoking rates or reductions in the number of teenage pregnancies. To paraphrase Lee (2013), doing good cannot be separated from doing right. To be morally grounded and accountable, the use of deception in health communication must be assessed using both teleological and deontological reasoning.

Although health communication is increasingly being held to higher moral standards, ethics still has not been accorded commensurate attention. Seedhouse (2004) observed that "ethics is rarely thought to be an issue in standard health promotion work" (p. 53). Unlike public relations, journalism, and advertising, the field of health communication has largely escaped scrutiny and criticisms primarily due to the noble intentions that underpin health messages. Health communication efforts are taken for granted to be benevolent endeavors grounded morally in noble, altruistic justifications and beneficent regard for others (Andreasen, 2001; Faden & Faden, 1978; Guttman, 1997, 2000, 2003; Kirklin, 2007a, 2007b; Kozlowski & O'Connor, 2003; Lee, 2011b; Lee & Cheng, 2010; Seedhouse, 1988, 2004). In its pursuit of a larger and highly tangible axiomatic good—individuals' and societies' well-being in health, health communication sets itself apart from other mass communicated, persuasive messages. This is particularly clear when compared to advertising and public relations that are seen as insidious activities seeking to alter individuals' attitudes and behaviors often to the detriment of their interests or well-being, with the goal of selling a product, service, or satisfying organizational goals (Jaksa & Pritchard, 1994). Health communication's values and motivations also attract little scrutiny and hence fewer criticisms due to the field's lack of commercial interests. The field also is protected by its historical associations with public service and the work of governmental agencies and charitable organizations, as well as what Guttman (2000) described as "a promise that scholarship, when applied to practice, can help individuals and groups with particular needs, or better society as a whole" (p. xii).

Few studies have examined truth-telling in public health messages although the use of deception has been directly and indirectly addressed in the context of health, specifically in physician-provided interactions (e.g., Fan & Li, 2004; Kirklin, 2007a; Kozlowski & Connor, 2003; Mattson & Roberts, 2001; Novack et al., 1989; Teasdale & Kent, 1995). Mattson and Roberts (2001), who studied the use of deception during HIV-test counseling, found that deception may be comfortably introduced into pre-HIV test counseling as a strategy, albeit controversial, for initiating healthful behaviors. When an individual is tested HIV-positive, the truth becomes an obstacle to the ability to initiate safer sex with their partners.

Although many consider public health messages to be a form of advertisement, albeit of a different scope and aim, there is comparatively less discussion in the literature about truth-telling in the context of public health messages. Kozlowski and Connor (2003) found that two respected agencies of the US Department of Health and Human Services (DHHS)—the Centers for Disease Control and Prevention (CDC) and the Substance Abuse and Mental Health Services Administration (SAMHSA)—maintained websites that misled readers into believing that smokeless tobacco is not safer than cigarettes. Kozlowski and Connor attributed such deception to "the desire to do

everything possible (including denying the truth and evading questions) to discourage the use of addictive smokeless tobacco" (p. 188). Although there is understandable reluctance to do anything that might encourage the use of addictive substances that pose health risks, the use of deception is highly troubling notwithstanding that "there will always be a gray area between truthful persuasiveness and outright deception" (p. 188).

Research on antismoking messages, one of the most widely studied bodies of health messages, also has focused on message efficacy and largely ignores the ethical dimensions. In a significant contribution, Lee (2011b) combined both frameworks—efficacy and ethicality—for understanding truth-telling. Based on a content analysis of 974 antismoking television PSAs, Lee (2011b) applied Baker and Martinson's (2001) TARES framework to examine how truthfulness was applied through eight dimensions and explored the relationships between message truthfulness and message attributes and audience characteristics. The study found that in general, antismoking PSAs reveal a high degree of truthfulness. There are significant relationships between message truthfulness and thematic frames, emotion appeals, source, age, social role and smoking status, and positive framing of consequences. Ads targeted at teens/youth and smokers tend to have lower message truthfulness than ads targeting older age groups and non-smokers. Ads with humor and fear appeals are found to be less truthful.

The rise of online health information seeking is rapidly transforming the landscape of health communication, with implications for the truth-telling imperative. As more individuals go online for health information, the changing power dynamics between health communicators and their audience would also have to address how facts are presented, accepted, and verified.

CONCLUSION

Deaver (1990) alluded to the importance of audience understanding of the fractional forms of truth in mass communication. Although Deaver did not elaborate, any framework of deception must consider the role of audiences and publics, who are no longer a passive, vulnerable, and persuadable lot as portrayed in the transmission model of communication. With the ubiquity of the Internet and new media technologies, the power dynamics between mass communicators and their audiences have changed considerably. It is increasingly difficult for mass communicators to deceive their audiences and publics. At the same time, it also is increasingly difficult for audiences and publics to determine what is deceptive and what is truthful. What is clear is that the moral imperative of truth underpinned by principles of beneficence and harm avoidance cannot change. Any departure from truth-telling, no matter how lofty the goal and how substantial the resultant social good, must be justified and explained by mass communicators, and be understood and accepted by audiences and publics.

References

Andreasen, A. R. (2001). *Ethics in social marketing.* Washington, DC: Georgetown University Press.

Armstrong, G. M., Gurol, M. N., & Russ, F. A. (1979). Detecting and correcting deceptive advertising. *Journal of Consumer Research, 6*(3), 237–246.

ASA Advertising Codes. (n.d.). Retrieved from https://www.asa.org.uk/codes-and-rulings/advertising-codes.html.

Bagdikian, B. (2000). *The media monopoly* (6th ed.). Boston: Preston Press.

Baker, S. (2008). The ethics of advocacy. In L. Wilkins & C. C. Christians (Eds.), *Handbook of mass media ethics* (pp. 115–129). Thousand Oaks, CA: Sage.

Baker, S., & Martinson, D. L. (2001). The TARES test: Five principles for ethical persuasion. *Journal of Mass Media Ethics, 16*(2), 148–176.

Barney, R. D., & Black, J. (1994). Ethics and professional persuasive communications. *Public Relations Review, 20*(3), 233–248.

Bivins, T. H. (2006). Responsibility and accountability. In K. Fitzpatrick & C. Bronstein (Eds.), *Ethics in public relations: Responsible advocacy* (pp. 19–38). Thousand Oaks, CA: Sage.

Bok, S. (1978). *Lying: Moral choice in public and private life.* New York: Pantheon Books.

Bonhoeffer, D. (1965). What is meant by telling the truth. In E. Bethge (Ed.), *Ethics* (pp. 363–372). New York: Macmillan.

Boush, D. M., Friestad, M., & Wright, P. (2015). *Deception in the marketplace: The psychology of deceptive persuasion and consumer self-protection.* New York and Hove: Routledge.

Bowen, S. A. (2006). *The business of truth: A guide to ethical communication.* San Francisco: IABC.

Bowen, S. A. (2008). A state of neglect: Public relations as corporate conscience or ethics counsel. *Journal of Public Relations Research, 20*(3), 271–296.

Broersma, M. (2010). The unbearable limitations of journalism on press critique and journalism's claim to truth. *International Communication Gazette, 72*(1), 21–33.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203–242.

Chisholm, R. M., & Feehan, T. D. (1977). The intent to deceive. *Journal of Philosophy, 74*(3), 143–159.

Code of Ethics for Professional Communicators. (2016). Retrieved from https://www.iabc.com/wp-content/uploads/2016/02/08-Code-of-Ethics_IEB-Approved_4-May-2016.pdf.

Curtin, P. A., & Boynton, L. A. (2001). Ethics in public relations: Theory and practice. In R. L. Heath & G. M. Vasquez (Eds.), *Handbook of public relations* (pp. 411–422). New York: Sage.

Darke, P. R., & Ritchie, R. J. (2007). The defensive consumer: Advertising deception, defensive processing, and distrust. *Journal of Marketing Research, 44*(1), 114–127.

Deaver, F. (1990). On defining truth. *Journal of Mass Media Ethics, 5*(3), 168–177.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995.

Edelman Code of Ethics and Business Conduct. (2017). Retrieved from https://www.edelman.com/about-us/code-ethics-and-business-conduct.

Elliot, D., & Culver, C. (1992). Defining and analyzing journalistic deception. *Journal of Mass Media Ethics, 7*(2), 69–84.

Faden, R., & Faden, A. I. (1978). The ethics of health education as public health policy. *Health Education & Behavior, 6*(2), 180–197.

Fan, R., & Li, B. (2004). Truth telling in medicine: The Confucian view. *Journal of Medical Philosophy, 29*, 179–193.

Fisher, C. (2016). Ten shades of truth: A study of Australian journalists' shift to political PR. *Public Relations Review, 42*(4), 665–672.

Fitzpatrick, K. R. (2002). Evolving standards in public relations: A historical examination of PRSA's codes of ethics. *Journal of Mass Media Ethics, 17*(2), 89–110.

Fitzpatrick, K., & Gauthier, C. (2001). Toward a professional responsibility theory of public relations ethics. *Journal of Mass Media Ethics, 16*(2–3), 193–212.

Fried, C. (1978). *Right and wrong.* Cambridge: Harvard University Press.

Gardner, D. M. (1975). Deception in advertising: A conceptual approach. *Journal of Marketing, 1*, 40–46.

Guttman, N. (1997). Ethical dilemmas in health campaigns. *Health Communication, 9*(2), 155–190.

Guttman, N. (2000). *Public health communication interventions: Values and ethical dilemmas.* Thousand Oaks, CA: Sage.

Guttman, N. (2003). Ethics in health communication interventions. In T. L. Thompson, A. M. Dorsey, K. I. Miller, & R. Parrott (Eds.), *Handbook of health communication* (pp. 651–679). Mahwah, NJ: Lawrence Erlbaum.

Hove, T. (2014). Ethical influence in health promotion: Some blind spots in the liberal approach. *Public Health Ethics, 7*(2), 134–143.

Hunt, J., & Manning, P. K. (1991). The social context of police lying. *Symbolic Interaction, 14*(1), 51–70.

Hyman, M. R., Tansey, R., & Clark, J. W. (1994). Research on advertising ethics: Past, present, and future. *Journal of Advertising, 23*(3), 5–15.

Ikonen, P., Luoma-aho, V., & Bowen, S. A. (2017). Transparency for sponsored content: Analysing codes of ethics in public relations, marketing, advertising and journalism. *International Journal of Strategic Communication, 11*(2), 165–178.

Jaksa, J. A., & Pritchard, M. S. (1994). *Communication ethics: Methods of analysis.* Belmont, CA: Wadsworth.

Kant, I. (1785/1993). *Grounding for the metaphysics of morals: On a supposed right to lie because of philanthropic concerns* (J. W. Ellington [1785/1993], Trans., 3rd ed.). Indianapolis, IN: Hackett.

Kent, M. L., & Taylor, M. (2002). Toward a dialogic theory of public relations. *Public Relations Review, 28*(1), 21–37.

Kieran, M. (1997). *Media ethics: A philosophical approach.* Westport, CT: Praeger.

Kirby, S. D., & Andreasen, A. R. (2001). Marketing ethics to social marketers: A segmentation approach. In A. R. Andreasen (Ed.), *Ethics in social marketing* (pp. 160–183). Washington, DC: Georgetown University Press.

Kirklin, D. (2007a). Truth telling, autonomy and the role of metaphor. *Journal of Medical Ethics, 33*(1), 11–14.

Kirklin, D. (2007b). Framing, truth telling and the problem with non-directive counseling. *Journal of Medical Ethics, 33*(2), 58–62.

Kozlowski, L. T., & O'Connor, R. J. (2003). Apply federal research rules on deception to misleading health information: An example on smokeless tobacco and cigarettes. *Public Health Reports, 118*(3), 187–192.

Lee, S. T. (2004). Lying to tell the truth: Journalists and the social context of deception. *Mass Communication and Society, 7*(1), 97–120.

Lee, S. T. (2011a). Conceptualizing ethical knowledge and knowledge transfer in public relations. *Public Relations Review, 37*(1), 96–98.

Lee, S. T. (2011b). Understanding truth in health communication. *Journal of Mass Media Ethics, 26*(4), 263–281.

Lee, S. T. (2013). Doing good, doing right: The ethics of health communication. In R. Ahmed & B. Bates (Eds.), *Health communication and mass media: An integrated approach to policy and practice* (pp. 195–210). London: Gower.

Lee, S. T., & Cheng, I.-H. (2010). Assessing the TARES as an ethical model for antismoking ads. *Journal of Health Communication, 15,* 55–75.

Lee, S. T., & Cheng, I.-H. (2011). Characteristics and dimensions of ethical leadership in public relations. *Journal of Public Relations Research, 23*(1), 46–74.

Lee, S. T., & Cheng, I.-H. (2012). Ethics management in public relations: Practitioner conceptualizations of ethical leadership, knowledge, training and compliance. *Journal of Mass Media Ethics, 27*(2), 80–96.

Lee, S. T., & Nguyen, H. L. (2013). Explicating the moral responsibility of the advertiser: TARES as an ethical model for fast food advertising targeting children. *Journal of Mass Media Ethics, 28*(4), 225–240.

Lentz, C. S. (1996). Fairness in broadcasting doctrine and the constitution: Forced one-stop shopping in the marketplace of ideas. *University of Illinois Law Review*, *271*, 1–39.

L'Etang, J. (1994). Public relations and corporate responsibility: Some issues arising. *Journal of Business Ethics, 13,* 111–123.

Mattson, M., & Roberts, F. (2001). Overcoming truth telling as an obstacle to initiating safer sex: Clients and health practitioners planning deception during HIV test counseling. *Journal of Health Communication, 13*(4), 343–362.

Mieth, D. (1997). The basic norm of truthfulness: Its ethical justification and universality. In C. Christians & M. Traber (Eds.), *Communication ethics and universal values* (pp. 87–104). Thousand Oaks, CA: Sage.

Novack, D. H., Detering, B. J., Arnold, R., Forrow, L., Ladinsky, M., & Pezzullo, J. C. (1989). Physicians' attitudes toward using deception to resolve difficult ethical problems. *JAMA, 261*(20), 2980–2985.

Parsons, P. J. (2016). *Ethics in public relations: A guide to best practice*. London: Kogan Page.

Patterson, P., & Wilkins, L. (2002). *Media ethics: Issues and cases* (5th ed.). Boston, MA: McGraw Hill.

Patterson, P., & Wilkins, L. (2013). *Media ethics: Issues and cases* (8th ed.). Boston, MA: McGraw-Hill.

Pearson, R. A. (1989). *A theory of public relations ethics, doctoral dissertation*. Athens, OH: Ohio University.

Petty, R. D. (2015). The historic development of modern US advertising regulation. *Journal of Historical Research in Marketing, 7*(4), 524–548.

PRSA Code of Ethics. (n.d.). Retrieved from https://www.prsa.org/ethics/code-of-ethics/.

Seedhouse, D. (1988). *Ethics: The heart of health care*. London: Wiley.

Seedhouse, D. (2004). *Health promotion: Philosophy, prejudice and practice*. Chichester: Wiley.

Seib, P., & Fitzpatrick, (1995). *Public relations ethics.* Belmont, CA: Wadsworth.

Singer, J. (2007). Contested autonomy: Professional and popular claims on journalistic norms. *Journalism Studies, 8*(1), 79–95.

Teasdale, K., & Kent, G. (1995). The use of deception in nursing. *Journal of Medical Ethics, 21*(1), 77–81.

Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology, 11*, 69–89.

Ullah, N., & Hussain, M. (2015). Impact of unethical advertising, misleading information or deceptive advertising on customer purchasing intention with mediating effect of word of mouth: Case of Pakistan. *International Journal of Innovation and Economic Development, 1*(4), 49–69.

Weaver, C. K., Motion, J., & Roper, J. (2006). From propaganda to discourse (and back again): Truth, power, the public interest and public relations. In J. L'Etang & M. Pieczka (Eds.), *Public relations: Critical debates and contemporary practice* (pp. 7–21). Mahwah: Lawrence Erlbaum.

Wilkins, S., Beckenuyte, C., & Butt, M. M. (2016). Consumers' behavioural intentions after experiencing deception or cognitive dissonance caused by deceptive packaging, package downsizing or slack filling. *European Journal of Marketing, 50*(1/2), 213–235.

Xie, G. X. (2016). Deceptive advertising and third-person perception: The interplay of generalized and specific suspicion. *Journal of Marketing Communications, 22*(5), 494–512.

Xie, G. X., Madrigal, R., & Boush, D. M. (2015). Disentangling the effects of perceived deception and anticipated harm on consumer responses to deceptive advertising. *Journal of Business Ethics, 129*(2), 281–293.

Deceptive Marketing Outcomes: A Model for Marketing Communications

Kim B. Serota

> Marketers are a special kind of liar. Marketers lie to consumers because consumers demand it. Marketers tell the stories, and consumers believe them. Some marketers do it well. Others are pretty bad at it. Sometimes the stories help people get more done, enjoy life more and even live longer. Other times, when the story isn't authentic, it can have significant side effects and consumers pay the price.
>
> —Seth Godin, *All Marketers are Liars* (2005)

It would be easy to begin by describing marketing deception as one more topic within the broad domain of human deception. Fundamentally, marketing is a human behavioral process, the facilitation of economic exchange that occurs between humans as producers and consumers. But, as Godin aptly notes, "marketers are a special kind of liar," and it would be wrong to limit the characterization of deceptive marketing without recognizing the uniqueness of the context. Marketing deception entails a set of parameters and antecedents that makes it both akin to and distinct from the way deception is viewed in the interpersonal communication context. To understand marketing deception, it is necessary to examine the ways in which marketing and, especially, marketing communications[1] differ from other forms of communication.

This chapter examines the parameters of marketing, looks at the evolution of deception as a component of the marketing disciple, discusses how

K. B. Serota (✉)
Department of Management and Marketing,
Oakland University, Rochester, MI, USA
e-mail: serota@oakland.edu

and why *outcomes* have become the locus of theoretical and practical marketing deception thinking, and concludes by proposing a theoretical model that brings marketing deception into the broader framework of mainstream theory and research on human deception. Theorizing on marketing deception is heavily laced with lists of deceptive claim types and debates over the nuances associated with interpreting and applying those typologies. Theory is also divided between a legal/regulatory perspective grounded in the rise of the Federal Trade Commission (FTC) and a cognitive/behavior perspective that has its roots in information processing and consumer behavior theories.

THE MARKETING CONTEXT FOR DECEPTION

The economic nature of marketing sets it apart from other forms of human interaction. That is not to discount the bargaining that occurs in day-to-day deception contexts. People trade information for information, friends negotiate as they plan their social interactions, and lovers swap sexual favors for socio-emotional benefits. But to a large extent, interpersonal exchange is ad hoc; it occurs in the moment as people go about structuring their daily thoughts and actions. As the study of behavioral economics demonstrates, everyday life is filled with trade-offs for which deception may be the chosen resolution strategy (Gneezy, 2005; Mazar & Ariely, 2006). However, marketing is not ad hoc; it is planned and executed, as are its many forms of deception.

Marketing Versus *Communication*

While communication is understood as purposive human behavior both for transferring information and for producing and reproducing shared meaning (Craig, 1999), the specific purpose of marketing is to create value. The American Marketing Association defines marketing from the producer perspective as, "the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large" (AMA, 2013). Marketing is the subset of business practices that affect consumers directly through exchange between buyer and seller. Consumer value, or utility, is created when the exchange satisfies the consumer's wants and needs. Producer value, or profit, is derived from receiving compensation. Marketing consists of a mix of activities involving four elements: product, price, promotion, and distribution (i.e., place). Products (goods and services), pricing, and the delivery of those products may be seen as symbolic transfers of meaning. For example, prices are often used to telegraph or signal ideas about the value, quality, and availability of the products to which they are attached. Marketing deception is, however, most often associated with advertising and promotions. Deceptions may occur anywhere in the marketing mix; however, an understanding of how and

why marketers deceive can be best achieved by examining communicative deception in marketing.

The Marketing Production Process

It should not seem extraordinary that, "the domain of marketplace deception is in many ways the polar opposite of the everyday lie-telling context" (Boush, Friestad, & Wright, 2009, p. 23). Effective and efficient personal communication that transfers information and seeks common understanding relies on the truth. Successful deception relies on the broad expectation that most communication is truthful (Levine, 2014). However, commercial communications have different criteria for success. For business, effective marketing communication is judged by conversion to sales not shared meaning, and efficiency is gauged by return on investment, not information transfer. Reducing the costs of marketing increases profits. If the truth is difficult to deliver, reshaping it or ignoring it altogether may be the less costly and more effective alternative.

Marketing is an organized business activity. Unlike typical human discourse, marketing involves planning, strategy, and formal execution. Explaining the nature of deceit found in marketing communications, Boush et al. (2009) continue:

> In marketplace deceptions, there is lying per se, and there is also a huge array of other clever, deceptive acts and tactics beyond a blatant lie. The deception agents are professionally trained and professionally invested in the success of their deceptions. They collaboratively plan a deceptive strategy, consider alternative combinations of tactics to accomplish it, pretest it, and revise it before using it on key targets, and then monitor and revise it once it begins. They use professional communication craftspeople to construct every element of it. They rehearse and rehearse until the speakers perfect their deliveries and the story presentation is as they intend it to be. They carefully assess targets' vulnerabilities, distinguishing the easy prey from the vigilant, skilled consumers. (pp. 23–24)

Marketing practice and marketing deception are part of a cultural production system that is intended to persuade consumers to make choices in favor of one product or brand over others being considered. The textbook egalitarian objective of marketing communications is engagement, "the planned, integrated and controlled interactive dialogues with key target audiences to help achieve mutually beneficial objectives" (Dahlen, Lange, & Terry, 2010, p. 3). The phrase "mutually beneficial" refers to value creation for *both* the consumer (in the form of satisfying wants and needs) and the producer (in the form of revenue and profits). While reciprocity is an ideal of the modern marketing concept, this altruistic view is at odds with the long-held belief that the sole purpose of a business is to generate profits for its shareholders

(Friedman, 1970). Even when businesses take the more contemporary view that sustainability depends on balancing shareholder value with stakeholder value (Laszlo & Zhexembayeva, 2011), structural factors in the buyer–seller relationship can still lead to deceptive practices. As Akerlof (1970) famously observed, the asymmetry of information available to the producer (marketer) affords a means to fulfill the profit maximization motive at the expense of the consumer. Businesses implicitly weigh the potential and perceived risks of deceptive messaging (i.e., impaired brand equity, loss of customers, and legal recourse) with the desired end goal of greater profitability. If a business deems the chance for increased profit to exceed the risk, it is more likely to engage in deceptive marketing practices. In effect, many marketers will give up some customer satisfaction for the opportunity to increase profits.

Marketing Is Legally Distinct from Other Communications

In commercial exchange, the free flow of information is viewed as essential to helping consumers know prices and make informed choices among products. Marketing deception occurs when a producer or seller uses deceptive practices to gain an unfair advantage over competitors or consumers by influencing or manipulating consumer choice. In the US, the Supreme Court ruled that only truthful commercial speech is protected; misleading and deceptive advertising and other related forms of promotion are not protected by the law.[2] This legal perspective is central to modern marketing practices and to theorizing about marketing deception. Hastak and Mazis (2011) contrast this with the dearth of psychological theory applied to marketing deception noting, "researchers have frequently focused on legal frameworks rather than on theory-driven predictions. These legal frameworks have typically been derived from an analysis of deceptive advertising case law" (p. 157).

The legal perspective is influenced most by US law. With the exceptions of the European Union, countries aligned with English law (e.g., Canada and Australia), and China, formal advertising law is limited. The international approach relies heavily on self-regulation (Petty, 2014). This may be a practical reflection; in 1996, nearly 40% of global advertising spending occurred in the US (Petty, 1997). In 2017, US ad spending was 35%; nearly two-thirds of global spending was concentrated in five countries: the US, China, Japan, the UK, and Germany (eMarketer, 2017).

EVOLUTION OF THE MARKETING DECEPTION CONCEPT

Deception as an element of marketing practice is as old as organized human trading behavior. The theoretical concept of advertising deception emerged (along with advertising) in the production era of the late 1900s. Today, marketing deception has diversified along with the systems used for communicating and delivering products from producers to consumers. From

bricks-and-mortar retailing to online marketplaces and from traditional newspaper advertising to the nuances of social media marketing, deception has transformed to fit the context.

Buyer Beware

Before the twentieth century, regard for commercial deceptive practices was embodied in the Latin phrase *caveat emptor*, meaning let the buyer beware. Product claims were often outrageous. Dr. Dyes Voltaic Belt promised complete restoration of vigor and manhood. Clark Stanley's Snake Oil Liniment pledged immediate relief for sciatica, toothache, and frost bites. *Caveat emptor* guided the consumer; the buyer was expected to be skeptical of claims, examine goods before purchasing, and make an informed decision. Sellers were under no legal obligation to assure that product claims were truthful or even to warrant that products would do what they purported to do. The legal system classified advertisements "as matters of opinion."

Antitrust and Deception

Commerce-related deception became an element of concern when the US government sought to protect businesses from each other. The Sherman Antitrust Act (1890) and Clayton Antitrust Act (1914) provided legislation to prevent anti-competitive practices that would increase costs to the consumer. Antitrust legislation briefly addressed the idea that commercial deception might create a competitive disadvantage, but it offered little in regard to protection for consumers. The concern was that lying to customers might result in one company stealing customers from another. Dishonest and misleading product and brand information, especially information about competing brands, might unfairly alter buyers' choices. Nineteenth-century advertising often portrayed competitors as imitators or ineffective alternatives. Brand names were frequently chosen to create confusion between start-ups and established brands. The FTC Act of 1914 charged the FTC with protecting companies from each other by outlawing unfair acts, practices, and methods that could affect commerce and limit competition. But as a goal, protecting consumers from lies and deceptions in advertising received only marginal consideration until the 1960s.

A First Legal Definition of Deception

Legally defined deceptive marketing emerged slowly. The FTC, created to enforce the Clayton Antitrust Act, provided legal guidelines including those against "deceptive acts or practices in or affecting commerce" (15 USC §45(a)(1)). The US Supreme Court in 1922 upheld an FTC interpretation that false advertising is a form of unfair competition. As marketing entered

the aggressive selling orientation of the mid-century, a subsequent 1931 ruling limited the ability of the FTC to protect consumers; the Supreme Court determined false advertising could not be prohibited unless there was evidence of injury *to a competitor*. Injury would become a defining characteristic for the concept of marketing deception.

The Wheeler-Lea Act (1938) amended the FTC Act by adding "unfair or deceptive acts or practices in commerce" to unfair methods of competition (Kintner, 1966). Giving the FTC injunctive power over false advertising of food, drugs, cosmetics, and therapeutic devices when that advertising was deemed to cause injury *to the public* provided the first significant emphasis on consumer injury. The Lanham Act of 1947 broadened the legal scope of deception. Misrepresenting competitors' products was also treated as deceptive marketing when claims would materially damage the competitor through the loss of sales or goodwill. However, there still was no universally accepted definition of deceptive marketing (Kottman, 1964).

During the 1950s, business attitudes shifted from the selling orientation to a marketing orientation, with greater emphasis on the wants, needs, and behaviors of consumers. In 1962, President Kennedy's consumer rights speech to Congress launched an era of consumerism in the US. Legislation such as the Fair Packaging and Labeling Act (1966) and the Truth in Lending Act (1969) were created to address deceptions directed to consumers. In 1972, the FTC Act was amended to extend its regulatory power to any product if false advertising represented a threat to public health or safety.

Outcome Emerges as a Key Criterion

Attempts to specify what was meant by false advertising remained elusive. Freer (1949) offered that, "good advertising not only tells the literal truth, but also avoids possible deception through subtle implication or omission" (p. 360). But despite expanded FTC powers, terms like *unfairness* and *deceptive* remained loosely defined. By the 1970s, scholarly influence began to impact advertising deception. Dillon (1973) attempted to formally define deceptive marketing by describing it as any discrepancy between how the product is portrayed in advertisements and its real-life performance. Aaker (1974) broadened this to *perceptual input*, stating deception was any difference between advertised and factual performance that caused detrimental consumer purchasing behavior. Dyer and Kuehl (1974) elevated the word *material* to represent the economic consequences when a marketer does not fully disclose all relevant information.

Researchers explored how consumers processed information and at what point deception actually occurs. Armstrong, Kendall, and Russ (1975, p. 235) suggested three necessary factors for a false marketing message to be classified as deceptive: (1) the consumer must perceive or notice the message; (2) the consumer must believe the (false) content of the message; and (3)

the message must be salient or impact the consumer's buying decision. All three criteria had to be met or the message was not considered deceptive, even if it contained an outright lie. Shimp and Preston (1981) supported this logic, stating real deception can only occur if consumer behavior is negatively affected; the presence of deception depends on its consequences. Not all researchers subscribed to the behavioral outcomes-based approach. Olson and Dover (1978) argued that deception happens sooner in the process, claiming "deception occurs when the consumer acquires demonstrably false beliefs as a function of exposure to the advertisement" (p. 30). By this definition, a marketing message should be classified as deceptive if it creates, exploits, or expands false belief regardless of what the consumer does (if anything) based on the false belief. Gardner (1975) also advocated defining deception based on consumers' perceptions, noting that deception exists if the consumer is left with an impression that is factual untrue.

In legal practice, the behavioral outcomes view prevailed. In 1977, the Supreme Court affirmed that false advertising is subject to restraint; the First Amendment right to commercial speech does not extend to misleading the consumer (*Bates v. State Bar of Arizona*, 1977). Following passage of the FTC Improvement Act of 1980, the FTC clarified that *consumer injury* is the central issue in cases alleging unfairness. A three-part test was established for consumer injury: (1) the injury must be substantial, (2) it must not be outweighed by countervailing benefits to consumers or competition, and (3) it must be an injury that consumers themselves could not reasonably have avoided. In the last part, the FTC meant deception—the withholding of information (omission) or the misrepresentation of information needed to make comparisons and reasoned purchase decisions.

The Federal Trade Commission Statement

In 1983, a prescriptive definition of deception, the "Policy Statement on Deception," was given by the FTC (1983) to the US House of Representatives. This statement would serve as a pivotal point in the legal and scholarly evolution of the marketing deception concept (Serota, 2014). The FTC statement included three key elements: (1) the misrepresentation, omission, or practice must be likely to mislead the consumer, (2) the practice must be considered from the perspective of a reasonable consumer, and (3) the deception must have materiality.

The 1983 FTC statement, which remains in current use, clarified a number of terms. According to the FTC, *misrepresentation* means a statement is contrary to fact. *Omission* is the failure to disclose qualifying information that would prevent a practice, claim, representation, or reasonable belief from being misleading. *Practices* are seller's actions; examples of misleading actions include applying false quality seals, bait and switch appeals, playing on consumer fears, or creating impressions that the consumer has won a prize or

is getting something for free. With regard to the second element, *reasonable consumer*, the FTC considers to whom the message is targeting and takes into account factors such as age, education, intellect, and frame of mind. For example, advertising targeted to children is evaluated differently from marketing appeals directed to typical adults. The third element brought outcomes into legal focus. A *material* misrepresentation means the information or practice is important to consumers and will affect their choices or actions with regard to a purchase decision. If the deception would not cause a reasonable consumer to alter their purchase decision, the misrepresentation is not considered material.

Material Injury

Under the FTC definition, whether or not outcomes are injurious determines if marketing has been deceptive. A consumer or buyer is considered injured when a misrepresentation, omission, or deceptive practice is material, regardless of intent. In the legal environment, material means that a consumer's actions or decisions are affected in ways that are different from what would have occurred had the consumer not been deceived. For example, aspirin is a specific pain-killing compound that is identical in all products called aspirin. If Brand A claims that its aspirin product is more effective than aspirin from Brand B and consumers then choose Brand A because of this claim, Brand B will suffer a loss in revenue. Thus, a material injury has occurred. If Brand C includes an additive with its aspirin to boost the pain-relieving strength but does not disclose the ingredient and its effects, and some consumers get sick because of this additive, a material injury has occurred. In both cases, resulting injury is the necessary and sufficient condition to establish that deceptive marketing has occurred.

   Under the law, an outcome of material injury determines when consumers have been deceived into acting to their own detriment. Further, this legal application of injury applies regardless of marketer intent. The ways in which marketing messages misrepresent, omit, or mislead are varied and are often situation-specific. Because of this, precedents of material injury arise from prior legal decisions and are applied to current or future legal proceedings. While the case law approach meets regulatory needs, it does little to help develop a theoretical view of what constitutes marketing deception.

MARKETING CLAIM TYPOLOGIES

Marketing scholars studying deception have developed generalized descriptions of marketing behaviors in order to define and identify when deception has occurred. Ideally, such definitions, usually in the form of typologies, would apply in a broad range of circumstances. However, consumer behavior theories grounded in psychology and theories of interpersonal

communication fail to account for the planned nature of marketing messages and are too broad or vague for legal application. Consequently, the majority of deception typologies have emerged from a legal/regulatory standpoint rather than the theory-driven perspective.

Marketing deception typologies range from complex and elaborate to compact and parsimonious. Gardner (1975) provided a simple taxonomy of three categories—unconscionable lie, claim–fact discrepancy, and claim–belief interaction—that builds on legal practice rather than consumer theory. Russo, Metcalf, and Stephens (1981) suggest a modified version that recasts Gardner's categories as fraud, falsity, and misleadingness. More recently, Xie and Boush (2011) succinctly summarized message types with the following categories: (1) falsity (can be verified objectively as untruthful), (2) omission (failure to disclose material consumption-related information), and (3) implication (which cannot be stated overtly, but misleads consumers through erroneous inferences about product or service attributes). Interwoven with the development of typologies has been the discussion of puffery. These are claims that have the characteristics of deception but are considered harmless and, therefore, from a legal perspective do not have a deceptive outcome. Finding the line between harmful and harmless deception has been a recurring theme among marketing deception scholars. Hastak and Mazis (2011) offer a theory-driven typology of truthful but misleading claims. Toward a proposed theory of marketing deception, this chapter presents two relevant approaches, the broad typology offered by Gardner and the focused *truthful but misleading* typology by Hastak and Mazis, as well as a separate discussion of puffery.

Gardner's Typology

Broad conceptualizations of marketing deception attempt to distinguish outright false messages from a middle ground that is neither truth nor lie but still negatively impacts the consumer. Gardner (1975) developed a three-part typology encompassing major characteristics of messages with a tendency to deceive. Gardner clearly noted that a useful definition of deception should focus on "the effect of the message on the consumer" rather than "the act of deceiving by the advertiser" (p. 42). The general form of Gardner's typology persists in many subsequent efforts to classify deceptive messages.

The Unconscionable Lie

When a marketer provides completely false information, the message is an unconscionable lie. The purpose is usually to deceive the consumer into making an incorrectly informed choice and is best equated with fraud. As Gardner explains, even if qualifying statements were added, "there would be no way for consumers to achieve the claimed benefits" (p. 42). Unlike claim–fact

discrepancy and claim–belief interactions, the unconscionable lie, or fraud, "assumes a deliberate attempt to create false beliefs about the product" (Russo et al., 1981, p. 120). Omission is a variant that may fall into this category when an advertiser intentionally excludes relevant information to alter consumer purchasing behavior.

Claim–Fact Discrepancy

When consumers encounter claims based on information that must be qualified in order to correspond to the ground truth, there is a claim–fact discrepancy. This can be a literal discrepancy or take other forms such as omission. Gardner offers the example of dandruff shampoo, which may work for people with one specific skin condition, but would not for the most common dandruff-causing problems. Claim–fact discrepancy is closely related to falsity, creating a false belief (Russo et al., 1981), and omission, which is leaving out some facts so that the resulting interpretation differs from reality (Rotfield & Rotzoll, 1980). With omission, the consumer may be exposed to literally true information, but is still deceived due to incompleteness. Preston (1977) refers to this as the expansion implication and the effect is "to imply some false widening or increasing of the value it expresses" (p. 159). Omission creates a claim–fact discrepancy even if the marketer did not realize erroneous inferences would result from these missing facts. These discrepant messages often use specific cues such as "inconspicuously qualified" statements (e.g., clarifying details are presented in smaller text than the literal claim so that consumers may not read them), reference prices (e.g., the marketer manipulates the advertised price, such as citing an inflated "suggested retail price," to optimize perceived consumer value), and comparative advertising with implied superiority claims. Failure to provide complete and unambiguous claims can lead to false impressions that do not align with the facts.

Claim–Belief Interaction

When the new information from an advertisement interacts with existing attitudes and beliefs, the result may be a deceptive belief about the product. In other words, the message creates, increases, or exploits a false belief about the product's expected performance without making either explicit or implied deceptive claims (Gardner, 1975; Pappalardo, 1997; Russo et al., 1981). Gardner offers this example: Suppose detergent manufacturers found that putting red and blue crystals in a product resulted in consumers attributing more cleaning power to those detergents, even if it wasn't true. By stating in an advertisement that the product has blue crystals (true), the ad would be deceptive because it promotes a superior cleaning power *belief* even though no claim for increased cleaning power was made. Gardner and Ross (1973) go so far as to suggest any advertising will be deceptive to some consumers because of the claim–belief interaction.

Puffery

While Gardner ([1975](#)) and others strived to identify claim types that would lead the consumer to a negative outcome, other research attempted to establish when consumers are *not* deceived. Advertising and other sales presentations often praise a product with vague and general subjective opinions, superlatives, or exaggerations, but without stating specific facts, this practice is known as "seller's talk," or *puffery* (Preston, [1977](#), [1996](#)). For example, when Starbucks advertises, "The best coffee for the best you," the *best coffee* claim is a puff, an exaggeration that cannot be proven true or false. Some consumers could be deceived if they believed the claim to be literally true but most consumers understand that the brand is just self-promoting.

This major category may or may not be a type of deceptive messaging, but it assumes certain claims are so obviously false that consumers won't rely on them (Rotfeld & Rotzoll, [1980](#), [1981](#)). Some scholars have asserted "puffery is common and expected" and, as a result, "consumers learn to discount… claims, thus protecting themselves from deception" (Urbany, Bearden, & Weilbaker, [1988](#), p. 95). Alternate results have been reported, especially in regard to referent prices. Kaufmann, Smith, and Ortmeyer ([1994](#)) found that inflated prices, resulting from puffery claims, can cause consumers to receive less value than expected. Thus, while some argue that puffery exists in the gray area between truth and lies, others argue that it is a form of deception.

The deceptiveness of puffery has been among the most perplexing of legal issues. Hoffman ([2006](#)), in an article with the intentionally ironic title, *The Best Puffery Article Ever*, presents both sides. On the one hand he says, "Speech that misleads consumers is presumptively unlawful" (p. 1400), but offers the frequently invoked counterargument, "It is puffery and should be immune from liability" (p. 1400). The puffery defense applies when claims are (a) incapable of measurement and (b) the consumer would not take them seriously.

Is Puffery a Misleading Inference?

Deceptive advertising and other marketing promotions can be explicitly false, or they may make misleading claims that are true but lead consumers into making false or incorrect inferences (Burke, DeSarbo, Oliver, & Robertson, [1988](#); Hastak & Mazis, [2011](#)). Consistent with claim–belief interaction, they may even provide truthful information with cultural referents or other implied meanings that lead the consumer to incorrect inferences about the product (Cohen, [2017](#)). Inferences of meaning are difficult to regulate. There may be considerable doubt about what has occurred in the mind of an individual consumer and whether or not the inference has led the consumer to a materially injurious outcome. Legally, puffery is often treated as allowable exaggeration, the FTC position being that consumers expect inflated claims in advertising, recognize it, and do not believe it. Oliver ([1979](#)) argues that this is a distinction without a difference; puffery is a claim made by

implication. However, the Uniform Commercial Code of 1996 sustained the right of advertisers to use puffery, placing the burden of proof that one has been misled on the buyer.

Marketing Deception Versus Everyday Deception

The idea of allowable exaggeration creates an important contrast between interpersonal and marketing deception. Truth-default theory (TDT) argues that most people are honest most of the time (Serota, Levine, & Boster, 2010) and are therefore truth-biased toward the messages they receive (Levine, 2014). Truth bias in turn makes receivers more susceptible to deception. Puffery exacerbates the consumer's difficulty with ascertaining the truth. To accommodate puffery, consumers as receivers lower their standard for what constitutes a truthful statement. This invites marketers (senders) to create incorrect inferences by using exaggerated and hyperbolic promotional marketing claims. If consumers (receivers) expect and treat less-than-fully-truthful information as a watered down truth, the value of the information to decision-making is also diluted, and the quality or outcome of the decision may be less certain than a decision made with full and truthful information. This, in turn, increases the probability of an injurious outcome.

Truthful But Misleading Typology

A discussion of puffery leads back to the question of when gray-area messages have deceptive outcomes. Among marketing deception scholars, a recurring issue is whether the legal or FTC perspective is over-adhered to while too little attention is given to attitude or behavioral theory (Hastak & Mazis, 2011). Legal theory is clear with regard to fraud (or the unconscionable lie) and, at least by case law precedent, provides for handling some falsity (or claim–fact discrepancies). However, the shadings of truth—claim–belief interactions, misrepresentations, and puffery—leave marketing scholars struggling to clarify deception, despite the efforts of regulators and legal scholars to give meaning to these labels. Hastak and Mazis propose a five-part typology of truthful but misleading claims, grounded in theory.

Omission of Material Facts

These occur when key facts have been left out of a claim. These can be complete omissions or half-truths, such as the "free" offer that does not disclose relevant terms of the offer. This type of claim is grounded in schema theory (Schank & Abelson, 1977) and Grice's (1989) theory of conversational norms (especially the maxim of quantity).

Misleadingness Due to Semantic Confusion

This type of claim occurs when the marketer uses unclear or deliberately confusing language, symbols, or images. For example, labeling processed foods as

"fresh frozen" creates confusion about the use of the word "fresh." The consumer may rightfully ask, "What does this mean?" This type of claim is based on the theory of and research on pragmatic implication (Harris & Monaco, 1978).

Intra-Attribute Misleadingness

This refers to a claim about an attribute that leads to a misleading inference about the *same* attribute. Typically, claims of uniqueness or performance fall into this category; for example, the claim that a product is gluten-free may lead to a false belief that competing products contain gluten when in fact they do not. In addition to the influence of pragmatic implication, this type of misleadingness is also found in research on "feature-absent" inferences (Burke, Milberg, & Moe, 1997).

Interattribute Misleadingness

In contrast to intra-attribute misleadingness, this type of claim occurs when the claim about one attribute leads to misleading inferences about *another* attribute. Consumers often associate multiple attributes with each other, although sometimes incorrectly. For example, consumers often conflate the dietary concerns of cholesterol and fat. When a claim is made that a product is low in cholesterol and the consumer interprets that to mean low in fat, interattribute misleadingness has occurred. Theoretically, this type of claim reflects inferences made through logical consistency. A number of studies demonstrate this kind of inference making (cf. Andrews, Netemeyer, & Burton, 1998). Hastak and Mazis further note that logical consistency can be exacerbated by belief preservation (Lord, Ross, & Lepper, 1979).

Source-Based Misleadingness

Advertisers frequently rely on the credibility of others to strengthen a claim. Source-based misleadingness occurs when an endorsement by expert or consumer testimonial is biased. Hastak and Mazis offer the examples of a surgeon endorsing a weight-loss supplement or an extreme weight-loss testimonial that doesn't reflect typical experience. Social influence theory (Deutsch & Gerard, 1955) suggests the tendency to accept information provided by other people as evidence of ground truth. Testimonials by consumers reflect social proof or the idea that if many people do or say something, it must be true (Asch, 1951; Cialdini, 2009).

THE PROBLEM WITH TRUTH IN MARKETING

An implication of the opening vignette is that what counts as truths and lies, and how the labels "truth" and "lie" are understood may vary by context. In everyday use, truth usually means a statement that has its basis in fact

while a lie is something that knowingly does not correspond to the ground truth. These same words have different meanings in marketing, an environment where messages are crafted to encourage and facilitate consumption. The outright lie—in marketing, a falsehood meant to deceive with the intent of enriching the deceiver to the detriment of the receiver—is considered an unethical lie, as it is in most other contexts. But the unblemished truth—the complete truth without omission or qualification—is problematic for marketers. In daily discourse, we accept a few white lies to avoid conflict or hurting someone's feelings. In marketing, the full truth often makes for a less than compelling story. Consumers rarely[3] encounter the claim, "Buy my product, it works … sometimes." In a marketing context, full truth can undermine the logic of a persuasive appeal.

An ethical conflict for marketers is that veracity is secondary to successful persuasion. Marketers do recognize the importance, and even the potential benefits, of ethical practices. However, the primary concern of the marketing function is to facilitate economic exchange. For marketers, exchange-promoting behaviors will supersede honesty as long as the exchange is in fact beneficial and does not injure the consumer. Many marketers assume that consumers understand this distinction. Marketers believe that unlike everyday conversation wherein there is a presumption of truth, the consumer has been conditioned to expect a persuasive appeal in marketing communications and will therefore approach all marketing messages with a dose of skepticism.

The subtlety of this exchange–veracity relationship is reflected in ideas such as *practical integrity*, which refers to imposing multiple forms of integrity *throughout* the marketing mix (Upshaw, 2007); in this view, integrity is embedded in the marketing process. Chonko and Hunt (1985, 2000) found that honesty is a top ethical issue for marketers but note the ethical conflict for marketing managers arises from having to *balance* the demands of the corporation against consumer needs. Thompson (2002) builds on the conflicted nature of marketing (consumerism versus ethical marketing) when proposing marketing virtues, among them honesty and safety. The virtue of veracity resides *within* the realm of marketing, not alongside or ahead of it. While making the cases against and for "shady practices," Jackson (1990) observes, "Provided everyone understands the rules of the advertiser's game, viz. that we have no right to trust their claims except so far as the law decrees, is it not permissible for them to gull and cajole us if they can into preferring their product over rivals?" (p. 58). Shading the truth may be a legal concern, but, beyond that, marketers are not compelled to undermine their own competitive efforts. Abela (2014) proposes the idea of "truths relevant to the audience" arguing that claims only need to be truthful, not explicitly true. In sum, truth may be a sought-after virtue for marketers, but it is also an obstacle.

A Model of Deception Marketing Outcomes

The marketing discipline struggles to coalesce around a unified theory of marketing deception. This is in part due to the entangled roots of the legal/regulatory perspective and the consumer behavior perspective. In this section, I propose a deceptive marketing outcomes (DMO) model for marketing deception, building on theory from communication and social psychology. This approach starts with a broad theory of deception, specifically Levine's (2014) truth-default theory (TDT), and recognizes several parameters that are specific to marketing (Serota & Levine, 2013). The proposed theory distinguishes deceptive from non-deceptive marketing and provides a structure for future research into marketing deception detection. It also suggests avenues that have been mostly ignored by the marketing discipline that could yield new insights into the messaging and cognitive processes by which consumers are deceived.

Theoretical Propositions

The DMO model assumes three main propositions that affect how we think about marketing deception, each of which is discussed in more depth below. First, the truth is less rigidly defined in marketing than in everyday discourse. Second, deception can occur with or without the sender's intent to deceive. Third, whether or not outcomes are injurious determines if *marketing deception* has occurred. The third proposition carries with it an important caveat. Even if the consumer is not materially injured (the legal definition), other outcomes may have consequences for how consumers process marketing messages, which in turn may increase or decrease the likelihood of being deceived. One intent of the model is to give a framework for integrating the existing body of research on marketing deception into the broader social-psychological deception theory framework.

Truth Is Not Rigidly Defined in Marketing
The first proposition of the DMO model reflects a key difference between marketing and cognitive science. It would appear that many marketers subscribe to a pragmatic theory of truth, in the mold of Charles Sanders Peirce or William James, rather than to the correspondence theory of truth originating with Aristotle ("what is that it is, is true") or the coherence theories of Spinoza and F. H. Bradley, and modern philosophers of science such as Carl Hempel and Karl Popper. For marketers, to paraphrase Pierce, the truth is that which is fated to be regardless of how we choose to get there (Houser & Kloesel, 1992). In a philosophy of marketing, outcomes determine *the truth*. The marketing message is truly deceptive only when the effect of the message is harmful.

Somewhat arbitrarily, I have chosen to refer to the kind of truth conceived in the world of commerce as *marketing-truth*. Gardner (1975) calls this *legal*

truth, but that may be too narrow given scholarly efforts to define marketing deception. Of the many paths an idea can take from sender to receiver to interpretation and outcome, as Fig. 42.1 illustrates, only when the marketer sends a deceptive message that (a) is judged by the consumer to be truthful and (b) is economically or otherwise injurious to the consumer will marketing deception occur. For the marketer, a factually correct message is truthful and therefore no deception has occurred. But an untruthful message is also considered not deceptive if the lie is recognized. If the untruthful message fails to deceive, the question of injury is moot. Messages that are recognized as untruthful but harmless are part of the marketing tableau; as such, these messages are treated as acceptable to the consumer; they are a marketing-constrained form of the truth, or marketing-truth. The possibility that there may be negative consequences along other paths (e.g., catching marketing lies could make consumers more suspicious of marketers generally) receives little consideration among marketing practitioners or scholars.

Marketing Deception Is Independent of Intent

The second DMO proposition indicates how profoundly the marketing perspective diverges from other views of deception. In most areas of deception theory, intent is a fundamental criterion for deception. In psychological and communication theories, there must be intent or motive in order for a message or an interaction to be deceptive. In the philosophy of deception, intent is a defining component. When Bok (1999) defines deception she writes, "When we undertake to deceive others intentionally, we communicate messages meant to mislead them, meant to make them believe what we ourselves do not believe" (p. 13). The only time that an exception is granted is in the area of psychiatry where deception and lying without intent are treated as pathological.

Perhaps marketing also deserves an exception. As discussed earlier in this chapter, marketing is a message production system. It consists of a series of actors—from the producers of the goods or services to be marketed to the people who craft and deliver the marketing messages through advertising, promotions, packaging, public relations, direct marketing, and personal selling. Products have attributes and qualities that are described in marketing messages and may include comparisons with competing products or qualifications for different audiences. And individual consumers each have existing attitudes and beliefs that can interact with the marketing message. While we might be able to trace the intent of a single producer's message through the process to an individual consumer, the diversity and quantity of products, messages, and receivers make it impossible to know whether any single marketing message transmission judged to be deceptive was deceptive by intent. Errors in production and transmission processes could account for the failure of an intentionally deceptive message to deceive. For example, many messages about diet supplements are intentionally deceptive because the producers are aware that the supplements offer little real health benefit (Fontanarosa,
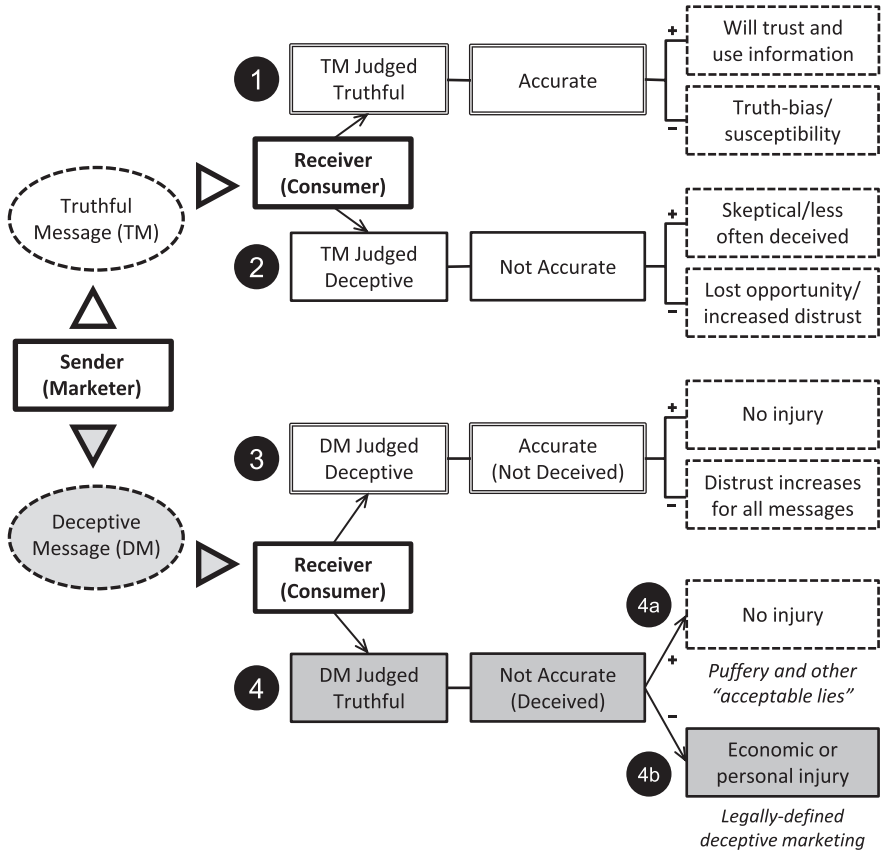
Fig. 42.1 Deceptive marketing outcomes model. The model shows four paths of messaging, truthful-deceptive judgments, and judgment outcomes. Each path has implications for deception theory. Cognitive theorists generally accept Paths 3 and 4 (with intent) as deceptive. Most marketing deception scholars accept Path 4 as potentially deceptive; the legal perspective only accepts Path 4b. Signs (+ and -) indicate positive and negative outcomes

Rennie, & DeAngelis, 2003). However, many consumers recognize these modern equivalents of snake oil and discount the category as well as the product, do not buy these products, and are therefore not deceived. Conversely, advertising for a product that has many qualities, only some of which get communicated, might be truthful and without intent to deceive. But if a feature that isn't communicated ends up injuring some consumers, the marketing has been deceptive. For example, a restaurant intends to satisfy customers with its cuisine but does not list all of the ingredients on its menu. While the restaurant may have had good intentions, when a customer with a specific food allergy to an undisclosed ingredient gets sick, the customer may be judged to have been deceived.

Material Injury Indicates Deception
The third DMO proposition draws on the legal/regulatory perspective; an outcome of financial injury to a competitor or personal injury (e.g., physical, emotional, or financial) to a consumer is sufficient for the marketing message to be deemed deceptive. Although the extent of injury might determine the outcome of individual legal proceedings, the FTC Statement of 1983 and subsequent interpretations are clear on this point. The important caveat for deception researchers is that legal deception and the processes by which consumers are deceived should be treated as separate issues. Marketing practitioners can draw from communication theory to develop a clearer conceptualization of how consumers are deceived. The model in Fig. 42.1 shows schematically the interaction between messaging, judgments, and alternative outcomes.

Four Message Paths to Deceptive and Non-Deceptive Outcomes

Communication theory (Bond & DePaulo, 2006; Burgoon & Levine, 2010) tends to evaluate deceptive messaging by considering four possible veracity conditions: (1) the message is truthful and the receiver judges it truthful, (2) the message is truthful and the receiver judges it deceptive, (3) the message is deceptive and the receiver judges it deceptive, and (4) the message is deceptive and the receiver judges it truthful. The DMO model describes each path from a marketing perspective, adding positive and negative outcomes to the underlying communication structure.

Path 1, the Message Is Truthful and the Receiver Judges It Truthful
This is the path of honest and accurate marketing. A marketer sends a truthful message (albeit one that satisfies the condition of marketing-truth), and the consumer recognizes it as truthful (even if it contains elements of normal exaggeration or puffery). The positive outcome is that the consumer will trust the information and use it to make a purchase decision that satisfies the wants and needs for which the product is intended. Consistent with Levine's TDT (2014), experiencing the truth tends to reinforce truth bias, a potentially negative outcome. Consumers who see mostly truthful messages expect truth to be the default for marketing messages and are, therefore, more susceptible to marketing deception.

Path 2, the Message Is Truthful and the Receiver Judges It Deceptive
This path describes what happens when the consumer is presented with the truth but perceives the message to be deceptive. In interpersonal deception, when there is a lack of transparency, a receiver may judge a truthful message as false. In marketing, this is most likely to occur when the marketing message is poorly executed. Do-it-yourself advertising or the use of incompetent marketing partners may result in badly crafted messages that are honest

but unreassuring. It may also occur when consumers are critical or have very high standards for truth. The positive outcome is that when consumers take this path they are more skeptical and less likely to be deceived. The negative outcome may be a lost opportunity to satisfy wants and needs appropriately (arguably a material injury). When consumers perceive they are being lied to, even when they are not, they may become less trusting and overly suspicious of all marketing messages or of marketing in general.

Path 3, the Message Is Deceptive and the Receiver Judges It Deceptive

When a consumer experiences this path, they recognize a deceptive message for what it is and are not deceived. This accurate judgment leads to an appropriate purchase decision (i.e., not buying), and consequently, the positive outcome is that no injury occurs. Armstrong et al. (1975) argue that if the consumer catches the deceptive message, no deception has occurred. However, communication theory would define the act of sending a deceptive message, especially if there is intent on the part of the sender, as deceptive. Olson and Dover (1978) align closely with this perspective, when they argue that exposure to deceptive marketing constitutes deception. It may be useful for marketing scholars to reconsider this view, as Path 3 may result in damage to the marketing process even if it is not injurious to the consumer. This is an important empirical question. Does exposure to deceptive messaging, even when accurately judged deceptive, increase consumer distrust for the marketing process? If yes, this would be a negative outcome of catching marketing deception.

Path 4, the Message Is Deceptive and the Receiver Judges It Truthful

This path is further divided based on the outcome. Path 4b is the path that is legally recognized as marketing deception when there is a negative outcome. On this path, a deceptive marketing message is misjudged by a consumer to be truthful, who then makes a purchase based on bad information, leading to material injury.

Path 4a, a deceptive message that is judged truthful but results in no injury, is problematic. If the consumer misjudges a deceptive message but is not injured, other factors must be moderating the outcome. Three possible situations might lead to this outcome. First, the consumer does not act even though they believe the marketing message. This might occur if the consumer intends to purchase but other influences such as competing needs, product unavailability, or insufficient resources keep the intention from being converted to a purchase (Ajzen, 1991). In this case, the deception is not validated. A second situation occurs when despite a successful deception, the consequences of purchasing are minor. For example, a consumer believes the claim that a brand of toothpaste will brighten teeth, but the outcome is that it does not work for that consumer. Since the purchase is inexpensive, easily replaced, and does no physical harm even though it also provides no

benefit, the injury is not material. The third situation arises when the message is deceptive but the individual consumer's standard for what is or is not truthful is very low. A consumer applying a very high standard for truth would be likely to judge the message deceptive (Path 3). But a consumer with a low standard might perceive the deceptive message to be puffery. As in the second situation, if the consumer purchases but the effects of purchasing are inconsequential, then the injury is not material. Communication theorists would find both Paths 4a and 4b to be deceptive, but legal-oriented marketing theorists are only likely to accept Path 4b as a deception.

Conclusion

This chapter has traced the concept of marketing deception from its legal/regulatory roots in antitrust legislation and the creation of the FTC to the present. Much of the scholarly research on marketing deception has informed the development of a legal test for deception in marketing communications. Despite changing marketing technologies, the outcome of processing a claim remains the central issue rather than how the claim is delivered or how deceptive claims are masked. Importantly, the effort to ground an understanding of marketing deception in consumer behavior or cognitive theories has been limited and is often in conflict with the legal perspective.

The proposed DMO model expands the theoretical approach to address sender–receiver consistency, similar to communication theory, and also adds outcomes that can occur in the marketing context. A key to reconciling marketing and communication theories is to recognize several elements that make marketing a distinct human activity. Marketing is a message production system; the goal of marketing activity is to facilitate economic exchange by informing, persuading, and reminding. Marketing primarily uses a one-to-many approach to communicate. In this context, marketers assume, and consumers are believed to accept, a lower standard for the truth (i.e., marketing-truth). Deception is viewed as occurring independent of intent. And, the concept of material injury is a defining criterion for judging whether or not consumers have been deceived.

The DMO model consists of the four combinations of sender messages (truthful or deceptive) by receiver judgments (truthful or deceptive). Each of the four paths has positive and negative outcomes. The DMO approach should help marketing deception scholars consider not only when deception has occurred (primarily in the deceptive message and truthful judgment condition), but also how outcomes in the non-deception conditions impact consumers' perceptions of marketing veracity and the marketing process overall. For communication theorists, the model suggests new avenues of deception research. In particular, the DMO model indicates a need to examine the cumulative effects of positive and negative outcomes on the formation of subsequent truth–lie judgments. It also invites comparisons of deception judgments in everyday discourse with judgments of structured and targeted

messages such as advertising, reinforcing Levine's (2014) emphasis on the importance of content in context.

NOTES

1. In the field of communication and other social sciences, the word "communication" is typically singular, referring to the transmission of information and the imparting of shared meaning. Among marketers and marketing scientists, "communications" is used as a specific reference to the multi-channel, multi-message nature of producer–consumer interactions; it is the act of creating messages in the promotions component of the marketing mix. In most instances, this text follows the convention of the discipline under discussion.

2. The First Amendment distinction is a notable source of confusion for consumers, particularly during the US American political season. Businesses and political campaigns share the vast array of public media channels in the US, and political campaigning has adopted many of the practices developed for marketing goods and services. However, commercial speech and political speech are regulated by two separate agencies of the government, guided by separate principles. Commercial speech is in the purview of the FTC, which has the authority to regulate and punish businesses that engage in deceptive communications. Political speech is regulated by the Federal Communication Commission (FCC), which is guided by the First Amendment. The FCC charge is to assure fair and equal access for all political candidates, but it has no authority to regulate the truthfulness or falsity of political campaign messages. Many media users may be unaware of this distinction. As a consequence, the uninformed voter who assumes all advertising is regulated may believe that political messages must be truthful in order to air. The cynical consumer who recognizes the dishonesty of many political ads may assume that commercial messages cannot be trusted either.

3. There are exceptions. For example, the pharmaceutical industry is bound by legal constraints to provide risk disclosures in their advertising and product information. As a consequence, drug advertising is frequently laden with warnings such as Product X "may cause heartburn, diarrhea, vomiting, and even death." Since these warnings are hardly purchase-inducing, advertising agencies try to moderate the negative effects by using techniques such as small type, accelerated reading, and weaving disclaimers into the story line of a commercial. The advertiser's intent is to keep the impact of these truths from undermining the persuasive product claims.

REFERENCES

Aaker, D. A. (1974). Deceptive advertising. In D. A. Aaker & G. S. Day (Eds.), *Consumerism: Search for the consumer interest* (pp. 137–156). New York: Simon & Schuster.

Abela, A. V. (2014). Appealing to the imagination: Effective and ethical marketing of religion. *Journal of Business Research, 67*, 50–58.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211.

Akerlof, G. (1970). The market for lemons. *Quarterly Journal of Economics, 84*(3), 488–500.

American Marketing Association. (2013). *Definition of marketing.* Retrieved from https://www.ama.org/AboutAMA/Pages/Definition-of-Marketing.aspx.

Andrews, J. C., Netemeyer, R. G., & Burton, S. (1998). Consumer generalization of nutrient content claims in advertising. *The Journal of Marketing, 62*(4), 62–75.

Armstrong, G. M., Kendall, C. L., & Russ, F. A. (1975). Applications of consumer information processing research to public policy issues. *Communication Research, 2*(3), 232–245.

Asch, S. E. (1951). Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177–190). Pittsburgh, PA: Carnegie Press.

Bates v. State Bar of Arizona. (1977). 433 U.S. 350.

Bok, S. (1999). *Lying: Moral choice in public and private life.* New York: Vintage.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214–234.

Boush, D. M., Friestad, M., & Wright, P. (2009). *Deception in the marketplace: The psychology of deceptive persuasion and consumer self-protection.* New York: Routledge.

Burgoon, J. K., & Levine, T. R. (2010). Advances in deception detection. In. S. Smith & S. Wilson (Eds.), *New directions in interpersonal communication* (pp. 201–220). Thousand Oaks, CA: Sage.

Burke, R. R., DeSarbo, W. S., Oliver, R. L., & Robertson, T. S. (1988). Deception by implication: An experimental investigation. *Journal of Consumer Research, 14*(4), 483–494.

Burke, S. J., Milberg, S. J., & Moe, W. W. (1997). Displaying common but previously neglected health claims on product labels: Understanding competitive advantages, deception, and education. *Journal of Public Policy & Marketing, 16*(2), 242–255.

Chonko, L. B., & Hunt, S. D. (1985). Ethics and marketing management: An empirical examination. *Journal of Business Research, 13*(4), 339–359.

Chonko, L. B., & Hunt, S. D. (2000). Ethics and marketing management: A retrospective and prospective commentary. *Journal of Business Research, 50*(3), 235–244.

Cialdini, R. B. (2009). *Influence, science, and practice* (5th ed.). Boston: Allyn & Bacon.

Cohen, S. (2017). Manipulation and deception. *Australasian Journal of Philosophy.* https://doi.org/10.1080/00048402.2017.1386692.

Craig, R. T. (1999). Communication theory as a field. *Communication Theory, 9*(2), 119–161.

Dahlen, M., Lange, F., & Smith, T. (2010). *Marketing communications: A brand narrative approach.* West Sussex: Wiley.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology, 51*(3), 629–636.

Dillon, T. (1973). What is deceptive advertising? *Journal of Advertising Research, 13*(5), 9–12.

Dyer, R. F., & Kuehl, P. G. (1974). The "corrective advertising" remedy of the FTC: An experimental evaluation. *The Journal of Marketing, 38*(1), 48–54.

eMarketer. (2017). *Worldwide ad spending: eMarketer's updated estimates and forecasts for 2016–2021* [Online executive summary]. Retrieved from https://www.emar-keter.com/Report/Worldwide-Ad-Spending-eMarketers-Updated-Estimates-Forecast-20162021/2002145.

Federal Trade Commission Act of 1914. (as amended 2006). Retrieved from https://www.ftc.gov/enforcement/statutes/federal-trade-commission-act.

Fontanarosa, P. B., Rennie, D., & DeAngelis, C. D. (2003). The need for regulation of dietary supplements—Lessons from ephedra. *The Journal of the American Medical Association, 289*(12), 1568–1570.

Freer, R. E. (1949). Informative and nondeceptive advertising. *Journal of Marketing, 13*(3), 358–363.

Friedman, M. (1970, September 13). The social responsibility of business is to increase its profits. *New York Times Magazine*.

FTC Policy Statement on Deception. (1983). Retrieved from http://www.ftc.gov/bcp/policystmt/ad-decept.htm.

Gardner, D. M. (1975). Deception in advertising: A conceptual approach. *Journal of Marketing, 39*, 40–46.

Gardner, D. M., & Ross, I. (1973, May). *Potential contributions of consumer psychology to deceptive advertising determinations and corrective measures*. Presented at 44th Annual Meeting of the Eastern Psychological Association, Washington, DC.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review, 95*(1), 384–394.

Godin, S. (2005). *All marketers are liars: The power of telling authentic stories in a low-trust world*. New York: Portfolio.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Harris, R. J., & Monaco, G. E. (1978). Psychology of pragmatic implication: Information processing between the lines. *Journal of Experimental Psychology: General, 107*(1), 1–22.

Hastak, M., & Mazis, M. B. (2011). Deception by implication: A typology of truthful but misleading advertising and labeling claims. *Journal of Public Policy & Marketing, 30*(2), 157–167.

Hoffman, D. A. (2006). The best puffery article ever. *Iowa Law Review, 91*, 1395–1448.

Houser, N., & Kloesel, C. (Eds.). (1992). *The essential Peirce: Volume 1: Selected philosophical writings (1867–1893)*. Bloomington: Indiana University Press.

Jackson, J. (1990). Honesty in marketing. *Journal of Applied Philosophy, 7*(1), 51–60.

Kaufmann, P. J., Smith, N. C., & Ortmeyer, G. K. (1994). Deception in retailer high-low pricing: A "rule of reason" approach. *Journal of Retailing, 70*(2), 115–138.

Kintner, E. W. (1966). Federal trade commission regulation of advertising. *Michigan Law Review, 64*(7), 1269–1284.

Kottman, E. J. (1964). A Semantic evaluation of misleading advertising. *Journal of Communication, 14*(3), 151–156.

Laszlo, C., & Zhexembayeva, N. (2011). *Embedded sustainability: The next big competitive advantage*. Stanford, CA: University Press.

Levine, T. R. (2014). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33*(4), 378–392.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098–2109.

Mazar, N., & Ariely, D. (2006). *Dishonesty in everyday life and its policy implications* (Working Paper Series, Federal Reserve Bank of Boston, No. 06-3). Boston, MA: Federal Reserve Bank of Boston.

Oliver, R. L. (1979). An interpretation of the attitudinal and behavioral effects of puffery. *Journal of Consumer Affairs, 13*(1), 8–27.

Olson, J. C., & Dover, P. A. (1978). Cognitive effects of deceptive advertising. *Journal of Marketing Research, 15*, 29–38.

Pappalardo, J. K. (1997). The role of consumer research in evaluating deception: An economist's perspective. *Antitrust Law Journal, 65*(3), 793–812.

Petty, R. D. (1997). Advertising law in the United States and European Union. *Journal of Public Policy & Marketing, 16*(1), 2–13.

Petty, R. D. (2014). International advertising law and regulation: A research review and agenda—The devil is in the details. In H. Cheng (Ed.), *The handbook of international advertising research* (pp. 395–410). Malden, MA: Wiley.

Preston, I. L. (1977). The FTC's handling of puffery and other selling claims made "by implication". *Journal of Business Research, 5*(2), 155–181.

Preston, I. L. (1996). *The great American blow-up: Puffery in advertising and selling*. Madison: University of Wisconsin Press.

Rotfeld, H. J., & Rotzoll, K. B. (1980). Is advertising puffery believed? *Journal of Advertising, 9*(3), 16–45.

Rotfeld, H. J., & Rotzoll, K. B. (1981). Puffery vs. fact claims—Really different? *Current Issues and Research in Advertising*, *4*(1), 85–103.

Russo, J. E., Metcalf, B. L., & Stephens, D. (1981). Identifying misleading advertising. *Journal of Consumer Research, 8*(2), 119–131.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Serota, K. B. (2014). Marketing, deceptive. In T. R. Levine (Ed.), *Encyclopedia of deception* (Vol. 2, pp. 643–647). Thousand Oaks, CA: Sage.

Serota, K. B., & Levine, T. R. (2013, July). *Truth bias and the detection of marketing deception*. Presentation at the 35th ISMS Marketing Science Conference, Istanbul, Turkey.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36*, 2–25.

Shimp, T. A., & Preston, I. L. (1981). Deceptive and nondeceptive consequences of evaluative advertising. *The Journal of Marketing, 45*(1), 22–32.

Thompson, M. (2002). Marketing virtue. *Business Ethics: A European Review, 11*(4), 354–362.

Upshaw, L. (2007). *Truth: The new rules for marketing in a skeptical world*. New York: AMACOM.

Urbany, J. E., Bearden, W. O., & Weilbaker, D. C. (1988). The effect of plausible and exaggerated reference prices on consumer perceptions and price search. *Journal of Consumer Research, 15*(1), 95–110.

Xie, G. X., & Boush, D. M. (2011). How susceptible are consumers to deceptive advertising claims? A retrospective look at the experimental research literature. *The Marketing Review, 11*(3), 293–314.

Audiences in the Dark: Deception in Pharmaceutical Advertising Through Verbal–Visual Mismatches

Viorela Dan

Consumer trust in advertising in general and pharmaceutical advertising in particular tends to be very low (Diehl, Mueller, & Terlutter, 2007; Steel, 1998). Indeed, advertising has a reputation for being deceptive (Faerber & Kreling, 2014; Shah, Holmes, & Desselle, 2003), and the actions of advertising practitioners are more often guided by financial considerations than ethical reasoning (Wilkins & Coleman, 2005). Recent surveys in the US found that advertising professionals are considered to be less honest and less guided by ethical standards than people in most other fields of activity (Gallup, 2012).

This chapter deals with pharmaceutical advertising, which seems to have a propensity to deception. Pharma ads appear inclined to praise the advertised drug exaggeratedly, overstate its health benefits, and make unsubstantiated claims linking therapy with the advertised drug to social rewards such as the love and acceptance of others (Abel, Neufeld, Sorel, & Weeks, 2008; Cline & Young, 2004; Jones, 1997; Messaris, 1997; Pallegedara & Hancock, 2014; Scalvini, 2010). This chapter starts from the observation that previous studies may have underestimated deception in pharmaceutical advertising by having focused on either verbal or visual means to deceive. Indeed, while verbal-only or visual-only studies have contributed tremendously to our understanding of deception in pharma ads, such an approach might no longer be suitable to

V. Dan (✉)
Ludwig Maximilian University of Munich, Munich, Germany
e-mail: viorela.dan@ifkw.lmu.de

deconstruct advertising practices which have gotten more and more sophisticated over the years. Newer studies suggest that modern-day deception is carried out by combining *conflicting verbal and visual claims* co-occurring in the same ad.

This has not been the focus of much research to date and is what this chapter deals with. This is justified with reference to the characteristics of most pharmaceutical ads: They are multimodal (i.e., consist of a verbal and a visual component). Thus, deception studies investigating words *or* visuals no longer suffice, as they can at most describe and explain only half of the message sent to audiences and thus only half of what carries the potential to deceive audiences. We delve into this topic after a brief introduction to pharmaceutical advertising. Research on deceptive pharmaceutical advertising is relevant to a wide range of fields of activity, including consumers, advertising regulators, drug companies, and advertising professionals.

Pharmaceutical Advertising

While other industries spend only 4–5% of their sales on marketing, pharma spends 20% (Dave & Kelly, 2014). Pharma marketing targets either health professionals or consumers (Greene & Herzberg, 2010; Wilkes, Bell, & Kravitz, 2000). Free samples and visits from pharma reps in the context of detailing together with ads and articles in medical journals are among the most common instruments for targeting health professionals. These are very costly undertakings to which 85% of the marketing budget is allocated (Baukus, 2004; Dave & Kelly, 2014).

The 15% spent by the pharma industry on consumer marketing may appear negligible. Yet, this amounts to almost $10 billion yearly in the US alone. Ad spending varies considerably by product category: dietary supplements (DS), over-the-counter or nonprescription drugs (OTC), and prescription drugs (Rx), respectively. It is estimated at over $900 million/year for DS, $3 billion/year for OTC, and at about $5 billion/year for Rx (DeLorme, Huh, Reid, & An, 2011). Every dollar invested in consumer drug ads returns many more in profit. Estimations range between $6 and over $20 in profit for each invested dollar (GAO, 2006; Hinshaw & Scheffler, 2014). For a long time, scholars believed that consumer advertising was able to increase sales for an entire class of products (e.g., antidepressants), but not necessarily for the specific drug advertised (e.g., Prozac) (Iizuka & Jin, 2007; Rosenthal, Berndt, Donohue, Epstein, & Frank, 2003). Newer studies generally confirmed that consumer ads raise class demand, but also found that they were successful in raising sales volume for the specific advertised drug. It is especially broadcast ads that seem able to increase both class demand and specific demand (e.g., Dave & Saffer, 2012).

DS are substances designed to enhance diet or correct nutritional deficiencies—as opposed to treating or curing illnesses. DS can be bought in

stores and pharmacies without prescription, i.e., they are unrestricted, as they are considered to lack abuse potential. OTC can too be bought in stores and pharmacies and do not require prescription. The difference is that OTC are intended to be used by consumers to self-treat mild illnesses. The abuse potential of OTC is nonexistent to low. Based on this, some OTC are restricted—i.e., handed out by pharmacists—while others are not. Finally, Rx drugs are developed to treat serious illnesses and can only be obtained with a prescription. Given the high abuse potential of many Rx, their use should be supervised by a physician (DeLorme et al., 2011).

Ads for all product categories are described as pervasive. Consumers turning on the TV are confronted to "a sea of consumer advertisements enticing them to eat burgers, buy cars, and talk to their doctors about drugs" (Coleman, 2010, p. 233). It is estimated that US TV viewers are exposed to over 30 hours of Rx ads per year (Brownfield, Bernhardt, Phan, Williams, & Parker, 2004). Perhaps due to the higher risks associated with Rx, most research to date focused on Rx ads—on product-claim ads especially.[1] Ads for DS and OTC remain under-researched (DeLorme et al., 2011).

A look at the breakdown of advertising spending by product category given above might raise the following question: Why does the industry spend such high amounts on Rx ads, when consumers are not able to purchase these drugs without a prescription? Indeed, unlike other types of ads, Rx ads do not simply have to convince consumers to buy the advertised product, as they are not allowed to do so. Rather, the purpose of Rx ads is to convince consumers to talk to their doctors about the advertised drug, i.e., to ask for a prescription (Dan, 2015, 2016). This call issued in Rx ads is answered by at least 30% of US Americans (KFF, 2001; Rodale Inc., 2003). The most advertised drugs are those that are selling best (Fisher & Ronald, 2008). This suggests that consumers' requests for prescriptions for the advertised drugs are met by their doctors. Currently, 70% of US Americans take at least one Rx drug (Zhong et al., 2013).

This mechanism of advertising influence can put pressure on doctors who may fear losing patients when denying them a prescription for the advertised drug. It is thus not surprising that health professionals have negative attitudes toward Rx ads. Some of the most common critiques regard deception. Many health professionals deem Rx ads to be misleading (Dan, 2015). When it comes to deception, ads for DS and OTC do not have a better reputation either (DeLorme et al., 2011). Accordingly, scholars and health professionals often call for stricter regulation of pharmaceutical advertising.

Deception in Pharmaceutical Advertising

This section begins with a brief introduction into the work of those institutions in charge of regulating pharmaceutical advertising in the US and the impact of various attempts to propose, loosen, or tighten regulations on

the amount and nature of pharmaceutical advertising. This provides a basis for addressing the second goal of this section, that of explaining deception accomplished by intentionally mismatching the verbal and the visual component of ads. This particular type of deception has been neglected in the literature so far.

A deceptive ad is one that makes claims that are misleading and/or not truthful, and that cannot be substantiated (An, 2014). More specifically, such an ad uses verbal, visual, or verbal and visual means to cause reasonable consumers "to acquire a false belief, or to continue to have a false belief, or to cease to have a true belief" (see Mahon, 2016, n.p.). In the case of pharmaceutical advertising, this could refer to the effectiveness, value, or features of the product. Consumers misled in such a way may be influenced to purchase the advertised drug or to ask their physicians for a prescription, if needed (An, 2014; Harris & Sanborn, 2009). They might needlessly adjust effective treatment regiments, be persuaded to switch a generic drug with a brand name, or over-medicate (Baukus, 2004; Carson, 2009). People may overbuy, self-medicate without need, or over-medicate (see Carson, 2009; Richards & Petty, 2007). Pharmaceutical advertising contributes to constantly pushing the boundaries of what constitutes a disease, to the effect that more and more natural bodily processes and mild or short-lived discomforts are transformed into medical problems (Carmack, 2014; Cline & Young, 2004; Kline, 2011; McAllister, 1992; Riggulsford, 2013). In this way, such ads help "create a drug culture in North America, where the solution to a problem is always swallowing a pill" (Coleman, 2010, p. 233).

Regulation

In the US, two institutions are in charge of regulating pharmaceutical advertising: the Federal Trade Commission (FTC) and the Food and Drug Administration (FDA). The FTC regulates ads for DS and OTC, while the FDA is in charge of Rx advertising. As the First Amendment also applies to commercial speech, only ads that do something illegal or misleading can be restricted (An, 2014). Barely any ads do something illegal (Macias, Lewis, & Tae Hyun, 2010), so regulators focus on deception. As already mentioned in the introduction to this chapter, previous content analyses suggest that pharma ads are often deceptive.[2]

The budgets of the FDA and the FTC, respectively, are far more modest than the advertising budget of the pharmaceutical industry (Baukus, 2004). This is highly consequential as the FDA can only review a very small sample of the ads submitted for review before dissemination (An, 2014; DeLorme et al., 2011). This is potentially worrisome, as consumers generally assume such reviews to be more expansive, and those who trust that authorities adequately regulate consumer advertising have a positive attitude toward it (Menon, Deshpande, Zinkhan, & Perri III, 2004).

The most important milestones in the regulation of pharma advertising were the following (An, 2014; DeLorme et al., 2011; Greene & Herzberg, 2010):

- 1985: The FDA makes the first attempt to regulate Rx ads. Ads were required to mention risks and benefits ("fair balance"), side effects and precautions, and a brief summary of warnings and precautions. This confined Rx ads to print, as fair balance could be satisfied here using small font size. Product-claim broadcast ads were virtually nonexistent; merely some help-seeking ads appeared on TV.
- 1994: The Dietary Supplement Health and Education Act (DSHEA) was passed in Congress. It was meant to increase access to and knowledge about DS.
- 1997: The FDA loosens regulations for Rx ads. It is now requested only that "adequate provision" for the dissemination of package labeling information is made (this replaced the "brief summary" stipulation). Adequate provision is met when consumers are referred to a hotline, a website, print ads, or health care professionals. Ads must now only address *major* risk and side effects.

DS advertising boomed after 1994, and Rx advertising boomed after 1997 (Baukus, 2004; DeLorme et al., 2011; Greene & Herzberg, 2010; Lee, 2010). The FTC standards for DS and OTC drug promotion require that claims "must be truthful, cannot be deceptive or unfair, and must be evidence-based" (FTC, 2016). Furthermore, as these ads make health claims, the FTC expects companies to be able to "support their advertising claims with solid proof" (FTC, 2016). For Rx ads, the FDA uses the term "substantial evidence" for a similar standard (An, 2014; DeLorme et al., 2011).

Verbal–Visual Mismatches

Like most messages in today's environment, pharma advertising is multimodal. This means that most ads consist of a verbal and a visual component. Print ads, as known from consumer magazines, are often two to three pages long. The first page consists of a full-page photograph and a slogan, while the other pages offer information on indication, risks, and side effects in small print. For instance, a recent ad for Isentress, a drug used in HIV therapy, shows two men cooking. The following statement is superimposed on the photograph "Hey Date Night! I love spending time with you. I was ready to learn more about my HIV treatment options. So I spoke to my health care professional and we chose ISENTRESS as part of my HIV regimen. He told me it could fight my HIV and may fit my needs and lifestyle. I can't wait to see you next time." The next two pages disclose several risks and side effects, ranging from dizziness to suicidal thoughts and actions.

TV ads are multimodal per definition. While they vary in length, they are often times constructed in a similar fashion. Ads begin by stating the health problem that the advertised drug hopes to tackle (i.e., the indication), continue with the benefits of starting treatment with the advertised drug (i.e., the health claims), and conclude with the mentioning of major risks and side effects. Sometimes, the health claim is repeated at the end, together with the appeal to talk to one's doctor about the advertised drug.

Most airtime/space in ads is dedicated to health claims regarding the drug's benefits. Moreover, benefits are prominently positioned, while risks are placed toward the end and given in small font and/or language that is either complex or vague (Dan, 2015). Risk information is typically addressed in the voice-over or the written text, while benefits are given in a multimodal format.

More importantly, verbal disclosures about a drug's risks and side effects are presented in a way that hinders their cognitive processing (Dan, 2015). Specifically, this negative verbal information is paired with positive visuals, which reiterate the health claims; show pleasant scenery; or depict healthy, athletic, and happy people enjoying themselves (Iyer, 2009). One study found that 91% of the analyzed Rx TV ads showed positive or neutral visual images during segments presenting risk information; none displayed negative visuals (Kaphingst, DeJong, Rudd, & Daltroy, 2004). By contrast, drug benefits are presented in a way that supports their cognitive processing (Dan, 2015): They are read out in segments showing congruent visuals (Iyer, 2009). In other words, the visuals fully illustrate or otherwise support the main health claims made in the ad (e.g., a person is shown sleeping in an ad, stating that the advertised drug helps people rest).

To illustrate, the table below offers a transcription of the voice-over and text on screen in a recent ad for Lunesta, a sleeping aid, together with a brief description of the accompanying visuals. Verbal–visual mismatches as described by Iyer (2009) and Kaphingst et al. (2004) and illustrated by the Lunesta ad are problematic (Table 43.1).

An older study by Wright (1979) tested if TV ads could motivate consumers to read product packages and read in-store warnings about OTC drugs. Ads were able to yield the desired effect when they showed people reading warnings/package information and the voice-over verbally instructed consumers to do so. Norris, Bailey, Bolls, and Wise (2012) showed that the recall of risk information was diminished when positive visuals were shown at the same time. Iyer and Feng (2011) tested the effect of modality dissonance during the presentation of risk information in ads for Rx drugs. These authors operationalized modality dissonance as the use of pleasant, non-risk visuals—e.g., images showing people running on the beach—while the voice-over addressed (serious) side effects of the advertised drug. The key finding was that the risk information in the voice-over was only recalled to a modest degree, as study participants processed the visuals instead of the voice-over. By contrast, study participants in the congruent condition recalled risk

Table 43.1 Description of a recent Rx ad (Available at https://www.youtube.com/watch?v=vu0rXFhsM8w)

	Verbal component		Text on screen	Visual component
	Voice-over			
Indication and Health Claim	We know a place where tossing and turning have given way to sleeping, where sleepless nights yield to restful sleep, and Lunesta can help you get there. Like it has for so many people before.		In clinical trials, Lunesta helped the majority of patients sleep up to 7 hours. Not actual patients, results may vary.	A woman is shown tossing and turning in her bed until a fluorescent butterfly flies through her bedroom window. The light dims and she falls asleep. The butterfly then enters another bedroom window, where a man is shown reading in his bed. He too falls asleep shortly after the butterfly entered his bedroom.
Risks and side effects	When taking Lunesta, don't drive or operate machinery until you feel fully awake. Walking, eating, driving, or engaging in other activities while asleep without remembering it the next day have been reported. Abnormal behaviors may include aggressiveness, agitation, hallucinations, or confusion. In depressed patients, worsening of depression, including risk of suicide, may occur. Alcohol may increase these risks. Allergic reactions such as tongue or throat swelling occur rarely and may be fatal. Side effects may include unpleasant taste, headache, dizziness, and morning drowsiness.		Take Lunesta before bed and only if you have 7 to 8 hours to sleep before becoming active. Lunesta has some risk of dependency. It is non-narcotic. Available by prescription only. See our ad in Food and Wine.	Viewers get a bird's-eye view of the city through which the butterfly continues its flight. The butterfly then enters a couple's bedroom; both are sound asleep.
Appeal	Ask your doctor if Lunesta is right for you and then get Lunesta for $0 at Lunesta.com.		Lunesta eszopiclone 1, 2, and 3 mg tablets. $0 co-pay at Lunesta.com Restrictions apply. Maximum reductions of $50 per prescription fill.	The butterfly exits the couple's bedroom and continues flying outside their house.
Health Claim	There is a land of restful sleep, we can help you go there, on the wings of Lunesta.		SUNOVION	LUNESTA.COM	1 (800) LUNESTA	Sunrise; the woman from the first shot wakes up rested and happy-looking.

information better (Iyer & Feng, 2011). Finally, part of the reason why the FDA found an advertisement for Strattera (an ADHD drug) to be deceptive was that it included "quick scene changes, visual changes, and erratic camera movement to minimize the display of side effects" (Hinshaw & Scheffler, 2014, p. 114). Russell, Swasy, Russell, and Engel (2017) also investigated the effects of mismatched words and visuals in pharma ads; for this, they used the term "cross-modality interference." These authors found that showing happy faces during a verbally delivered health warning reduces audiences' understanding of the warning (Russell et al., 2017).

Overall, audiences understand and remember information on risks and side effects to a lesser degree than on benefits (Kaphingst, Rudd, DeJong, & Daltroy, 2005; Menon, Deshpande, Perri III, & Zinkhan, 2003). Because words and visuals are only congruent in the segments in which benefits are presented, information on risks and side effects in consumer ads are often deemed difficult to understand and remember (Kaphingst et al., 2005).

At least nine theories and models have been proposed to explain differences in the cognitive processing, memory and recall of congruent multimodal messages and that of mismatched verbal–visual messages. While many of these theories and models were developed with news in mind, it seems plausible to assume that the underlying mechanisms they describe equally apply to other forms of storytelling, advertising included.

Almost 50 years ago, Posner, Nissen, and Klein (1976) argued the existence of *Visual Dominance*, which meant that visuals demand attention and high resources for processing. This was expected to lead to a situation where only limited resources were available for information presented through other modalities. A decade later, Paivio (1986) proposed the *Dual Coding Theory* (DCT). This theory started from the premise that the human brain has two subsystems at its disposal: one for the processing of words, and another for the processing of visuals. Information processing was described as variably fast and deliberate in each subsystem. The verbal subsystem is slower because processing is deliberate and sequential. The visual subsystem, by contrast, is fast due to the capacity to process visuals automatically and straightaway. Paivio (1986) explained further that words are stored only in the verbal subsystem (i.e., just once), whereas visuals are stored in the verbal and the visual subsystem (i.e., twice). This makes visuals easier to retrieve from memory. When words and visuals are used together, congruence is likely to improve memory and recall, just as incongruence (i.e., mismatched verbal and visual components of a message) is likely to hinder memory and recall. Paivio (1991) stated that "verbal and nonverbal codes corresponding to the same object (e.g., pictures and their names) can have additive effects on recall" (p. 259). By contrast, when modalities are mismatched, people tend to (mis-) remember information according to the visuals, something known as the *Picture Superiority Effect* (PSE).

The idea that words are processed analytically and visuals superficially was also at the core of two other models: Chen and Chaiken's (1999) *Heuristic Systematic Model* (HSM) and Cacioppo and Petty's (1984) *Elaboration Likelihood Model* (ELM). At least five other theories and models have taken on the notion that visuals are processed ahead of words and also faster than words. The *Cue-Summation Theory* (Reese, 1984) explained that modality congruence helps processing and recall because one modality offers cues for the other. By contrast, incongruence hinders learning due to unmatched cues. The *Belongingness Hypothesis* (Grimes, 1991) elaborated on why congruence helps information processing and incongruence hinders it. This is because conflicting modalities are "perceived as a semantic unit…because attention does not have to be distributed among different stimuli" (Grimes, 1991, p. 270). When modalities "are discordant enough, they will be regarded as separate units, each demanding attention" (Grimes, 1991, p. 271). The *Multiple Resource Theory* (Basil, 1994) argued similarly that messages in a video format prompt a competition between the modalities for resources. Modality congruence helps processing, otherwise visuals overpower the verbal stream. By contrast, incongruence is detrimental; even written text paired with visuals is overtaxing, because they both require visual resources.

In more recent years, the *Limited Capacity Model* (LCM) and the *Semantic Overlap Hypothesis* were proposed. The former was introduced by Lang (2000) and postulated that information processing is strained by multimodal video formats. The implication was that, in complex messages, visuals are processed while the other modalities are ignored. The latter was articulated by Walma van der Molen and Van der Voort (2000). These authors found that video messages with congruent modalities are processed and recalled easier than print. When attention is overtaxed, priority is given to the visual.

All these theories and models point in the same direction. Taken together, they clearly suggest that illustrating a verbal message with visuals telling a different story will diminish memory and recall of what is said. By contrast, adding congruent visuals to a verbal message improves memory and recall. One can thus hypothesize that ad segments that use congruent verbal and visual information are likely to be remembered, while viewers will likely only remember the visuals in segments using incongruent modalities. To the extent that the congruent modalities are used for the presentation of drug benefits, while incongruent modalities are used for risk and side effects information, FDA requests for a fair balance between drug benefits and risks/side effects is violated. Viewers are hindered in their capacity to absorb information about risks and side effects. There is reason to believe that verbal–visual mismatches are purposefully employed to distract audiences from disclosures on drug risks and side effects. In the words of Aditya (2001), advertisers are aware of "various psychological processes that limit a consumer's ability to respond in self-defense under the circumstances, such as the state of arousal

brought about by visual and verbal appeals that make some product features salient and others inconspicuous" (p. 747).

Discussion and Conclusion

This chapter started from the observation that even though most pharma ads are multimodal (i.e., have both a verbal and a visual component), most research to date focuses almost exclusively on the verbal component. Pharmaceutical products can be advertised only when certain principles are taken into account. One of the most important principles is that ads entail a fair balance between a drug's risks and benefits. From this perspective, a large body of research has focused on the verbal component of ads and addressed questions pertaining to compliance with this very principle of fair balance (e.g., if an appropriate amount of time or space was allotted to the presentation of drug benefits and drug risks, respectively). While these studies yielded very interesting results, they disregarded the visual component of ads. Ignoring half of the message sent to audiences suggests that only half of the effects are currently understood.

This chapter argued in favor of investigating deception in pharmaceutical advertising by looking at how verbal elements of the message are combined with visual elements. To date, regulation does not address this aspect and there is reason to believe that change is needed in this regard. While regulators tend to focus on words, advertisers have learned to use visuals and verbal–visual mismatches to increase the chances of yielding the desired effect despite regulation (Richards & Petty, 2007). This is problematic because of the potential for deception.

The purpose of this chapter was to offer a brief review of literature on the prevalence of verbal–visual mismatches in pharmaceutical advertising and their effects. It can be concluded that such mismatches are used strategically to distract audiences from risk information. Surely, pharma advertising remains advertising and its main purpose is to boost sales, regardless of the more often stated goal of educating consumers and improving public health (Dan, 2016). However, these ads are regulated for a reason, and it appears that advertisers take advantage of the weaknesses of the human brain in information processing. In other words, they appear to follow the letter of the law, but not the spirit of the law. Thus, while the strategic use of verbal–visual mismatches is not prohibited or illegal, it is certainly deceptive.

Potential consumers want to be informed about the risks and side effects of the drugs they consider taking (Everett, 1991). This information has a major influence in deciding if one wants to start treatment with that drug or not (Callaghan, Laraway, Snycerski, & McGee, 2013; Polen, Khanfar, & Clauson, 2009). Using pleasant imagery when this type of information is narrated distracts attention from unpleasant information about risks and side effects. Such verbal–visual mismatches are thus against the spirit of regulations currently in place and at odds with what consumers deem desirable.

Despite this criticism, one should resist the temptation to vilify pharma advertising across-the-board. If deception were the norm, advertising would be unable to increase sales and thus senseless. In Carson's (2009) words, it seems unlikely that deceitful advertisers would want their peers in the field to follow suit; they merely "want to make a special exception for themselves" (p. 254).

Indeed, the evidence gathered until now on verbal–visual mismatches during segments on risks and side effects is quite compelling. Should these findings be replicated in large-scale content analyses of pharma ads, then this would have to be addressed by regulators. Such content analyses could assist regulators by offering an assessment of the extent of the problem and a categorization of the strategies used. The next step would be to conduct experimental studies to understand the effects of deception through verbal–visual mismatches on the audience (see also Drumwright, 2007; Richards, 1990). It would then be up to the FDA, the FTC, and similar organizations worldwide to determine if stricter regulation is necessary.

Previous studies suggest that verbal–visual mismatches are pervasive during the presentation of risk information (e.g., images of happy-looking people are shown while serious risks are read aloud). Furthermore, it seems reasonable to assume that this hinders audiences in taking in risk information. Should this hold true, then regulators would have to address this. Yet, any attempt at forcing advertisers to illustrate segments presenting risks and side effects with actual depictions is likely to fail. In addition, it is questionable that this would lead to the desired effect. This is because drugs' side effects can be very serious or at least unaesthetic. Examples include sickness, rashes, and suicidal thoughts. Accordingly, it is hard to imagine that pharmaceutical companies would be willing to illustrate statements like "risk of suicide may occur" with congruent visuals. Even if they would, the result would be very disturbing and arguably counterproductive from a public health perspective.

In search for more realistic solutions, regulations could stipulate the use of visuals that are not dissonant during the presentation of risk information. Iyer and Feng (2011) suggested here talking heads (e.g., a doctor talking into the camera) and text boxes shown on screen. But previous studies suggest that talking heads can be distracting too (Dan, 2018), so the solution may involve using text boxes and narrating them in the voice-over. Future research could test several other options that would be acceptable to both public health officials and pharmaceutical companies.

From a political point of view, it appears sensible to increase the budget of the regulatory bodies, mainly the FTC and the FDA, to ensure premarket reviews of a larger sample of pharma ads. Perhaps workshops could be offered to the FDA/FTC employees actually conducting these reviews in order to help them identify such subtle forms of deception. With larger budgets, regulators could even make more use of consumer research (e.g., experiments) to help determine the extent of deception. Finally, it appears that deceptive ads should be sanctioned consistently. Regulators and advertisers can agree to

stop disseminating specific ads deemed deceptive by signing consent decrees. Further, regulators can issue cease-and-desist orders, demand that consumers are repaid, demand affirmative disclosure, enforce corrective advertising, or even pursue litigation (An, 2014; Richards & Petty, 2007). However, as corrective advertising may not yield the desired effect (Johar, 1996; Kollath-Cattano et al., 2014), regulators may be best advised to focus on preventing deceptive advertising rather than attempting to correct the harm done. Regulators' focus should be expanded to include verbal–visual mismatches, as modern-day deception in pharmaceutical advertising no longer occurs through verbal means or visual means alone. After all, pharma advertising is allowed under the condition that a fair balance between risk and benefit information is provided. When verbal–visual mismatches are used in segments disclosing risks and side effects, fair balance cannot be reached.

Notes

1. Product-claim ads are what scholars typically refer to when using the term direct-to-consumer pharmaceutical advertising, abbreviated DTCA. Product-claim ads name the drug brand and its purpose, as well as the drug's benefits and risks. Two other, less-controversial, types of Rx ads exist: reminder ads and help-seeking ads. Reminder ads show the drug brand, but not its purpose, benefits, and risks. Help-seeking ads inform about a disease or medical condition without naming drug brand. Product-claim ads are only allowed in US and New Zealand. Sustained lobby efforts are made to change regulation in the European Union (EU) (Dan, 2015).
2. Some additional examples for visual deception include the way pharma ads instruct viewers to link the way ad protagonists look or feel with taking the product advertised: When a very muscular person is shown in a DS ad for muscle growth, the implication is that this DS causes muscle growth (Messaris, 1997). This is also the case with Rx ads juxtaposing images of people looking distraught and then—after starting treatment with the advertised drug—people looking happy and having a good time with romantic partners or family members (Cline & Young, 2004). These ads would be deceptive if advertisers were unable to provide "solid proof" or "substantial evidence" for these claims.

References

Abel, G. A., Neufeld, E. J., Sorel, M., & Weeks, J. C. (2008). Direct-to-consumer advertising for bleeding disorders: A content analysis and expert evaluation of advertising claims. *Journal of Thrombosis and Haemostasis, 6*(10), 1680–1684. https://doi.org/10.1111/j.1538-7836.2008.03083.x.

Aditya, R. N. (2001). The psychology of deception in marketing: A conceptual framework for research and practice. *Psychology and Marketing, 18*(7), 735–761. https://doi.org/10.1002/mar.1028.

An, S. (2014). Advertising, governmental regulation of. In T. L. Thompson (Ed.), *Encyclopedia of health communication* (pp. 36–38). Thousand Oaks, CA: Sage.

Basil, M. D. (1994). Multiple resource theory I: Application to television viewing. *Communication Research, 21*(2), 177–207. https://doi.org/10.1177/009365094021002003.

Baukus, R. (2004). DTC advertising. *Journal of Health Communication, 9*(6), 563–564. https://doi.org/10.1080/10810730490882775.

Brownfield, E. D., Bernhardt, J. M., Phan, J. L., Williams, M. V., & Parker, R. M. (2004). Direct-to-consumer drug advertisements on network television: An exploration of quantity, frequency, and placement. *Journal of Health Communication, 9*(6), 491–497.

Cacioppo, J. T., & Petty, R. E. (1984). The elaboration likelihood model of persuasion. *Advances in Consumer Research, 19,* 673–675.

Callaghan, G. M., Laraway, S., Snycerski, S., & McGee, S. C. (2013). Antidepressant advertising effects on drug knowledge and drug seeking. *Journal of Consumer Marketing, 30*(3), 267–272. https://doi.org/10.1108/07363761311328937.

Carmack, H. J. (2014). Medicalization. In T. L. Thompson (Ed.), *Encyclopedia of health communication* (pp. 845–846). Thousand Oaks, CA: Sage.

Carson, T. L. (2009). Deception and information disclosure in business and professional ethics. In G. G. Brenkert (Ed.), *The Oxford handbook of business ethics*. Oxford: Oxford University Press.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chen & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). New York, NY: Guilford.

Cline, R. J. W., & Young, H. N. (2004). Marketing drugs, marketing health care relationships: A content analysis of visual cues in direct-to-consumer prescription drug advertising. *Health Communication, 16*(2), 131–157.

Coleman, C.-L. (2010). Drug advertising. In S. H. Priest (Ed.), *Encyclopedia of science and technology communication* (pp. 233–236). Thousand Oaks, CA: Sage.

Dan, V. (2015). Patientengerichtete Werbung für verschreibungspflichtige Medikamente (DTCA)—Überblick und Forschungslücken [Direct-to-consumer Advertising for Pharmaceutical Products (DTCA)—A literature review and directions for future research]. In M. Schäfer, O. Quiring, C. Rossmann, M. Hastall, & E. Baumann (Eds.), *Gesundheitskommunikation im gesellschaftlichen Wandel: Chancen und Herausforderungen* [Health communication and social change: Opportunities and challenges] (pp. 63–73). Baden-Baden, Germany: Nomos.

Dan, V. (2016). Pharmakommunikation [On the Strategic Communication of Pharmaceutical Companies]. In C. Rossmann & M. R. Hastall (Eds.), *Gesundheitskommunikation: Kommunikationswissenschaftliche Grundlagen* [Health Communication: Communication Basics]. Wiesbaden, Germany: Springer VS.

Dan, V. (2018). A methodological approach for integrative framing analysis of television news. In P. D'Angelo (Ed.), *Doing news framing analysis* (pp. 191–220). New York, NY: Routledge.

Dave, D. M., & Kelly, I. R. (2014). Advertising as a determinant of health in the USA. In A. J. Culyer (Ed.), *Encyclopedia of health economics* (pp. 32–50). Amsterdam: Elsevier.

Dave, D. M., & Saffer, H. (2012). Impact of direct-to-consumer advertising on pharmaceutical prices and demand. *Southern Economic Journal, 79*(1), 97–126.

DeLorme, D. E., Huh, J., Reid, L. N., & An, S. (2011). Advertising in health communication. Promoting pharmaceuticals and dietary supplements to U.S. consumers. In T. L. Thompson, R. Parrott, & J. F. Nussbaum (Eds.), *The Routledge handbook of health communication* (pp. 268–290). New York, NY: Routledge.

Diehl, S., Mueller, B., & Terlutter, R. (2007). Skepticism toward pharmaceutical advertising in the U.S. and Germany. *Cross-Cultural Buyer Behavior: Advances in International Marketing*, *18*, 31–60.

Drumwright, M. E. (2007). Advertising ethics: A multi-level theory approach. In G. J. Tellis & T. Ambler (Eds.), *The Sage handbook of advertising* (pp. 398–415). Thousand Oaks, CA: Sage.

Everett, S. E. (1991). Lay audience response to prescription drug advertising. *Journal of Advertising Research, 31*(2), 43–49.

Faerber, A., & Kreling, D. (2014). Content analysis of false and misleading claims in television advertising for prescription and nonprescription drugs. *JGIM. Journal of General Internal Medicine, 29*(1), 110–118. https://doi.org/10.1007/s11606-013-2604-0.

Fisher, J. A., & Ronald, L. M. (2008). Direct to consumer responsibility: Medical neoliberalism in pharmaceutical advertising and drug development. *Patients, Consumers and Civil Society, 10,* 29–51.

Federal Trade Commission (FTC). (2016). Advertising and marketing—Health claims. Retrieved from https://www.ftc.gov/tips-advice/business-center/advertising-and-marketing.

Gallup. (2012). Congress retains low honesty rating. Retrieved from http://news.gallup.com/poll/159035/congress-retains-low-honesty-rating.aspx.

Government Accountability Office (GAO). (2006). *Prescription drugs: Improvements needed in FDA's oversight of direct-to-consumer advertising*. Washington, DC: Author.

Greene, J. A., & Herzberg, D. (2010). Hidden in plain sight. Marketing prescription drugs to consumers in the twentieth century. *American Journal of Public Health, 100*(5), 793–803. https://doi.org/10.2105/ajph.2009.181255.

Grimes, T. (1991). Mild auditory-visual dissonance in television news may exceed viewer attentional capacity. *Human Communication Research, 18*(2), 268–298.

Harris, R. J., & Sanborn, F. W. (2009). *A cognitive psychology of mass communication* (6th ed.). London: Routledge.

Hinshaw, S. P., & Scheffler, R. M. (2014). *The ADHD Explosion: Myths, medication, money, and today's push for performance*. Oxford: Oxford University Press.

Iizuka, T., & Jin, G. Z. (2007). Direct to consumer advertising and prescription choice. *The Journal of Industrial Economics, 55*(4), 771. https://doi.org/10.1111/j.1467-6451.2007.00329.x.

Iyer, N. (2009). *This drug's for you: A content analysis of televised drug commercials* (PhD), Indiana University, Bloomington, IN.

Iyer, N., & Feng, Y. (2011). *Effect of information redundancy on disclosure recall in DTC commercials*. Paper presented at the annual meeting of the International Communication Association (ICA), Boston, MA.

Johar, G. V. (1996). Intended and unintended effects of corrective advertising on beliefs and evaluations: An exploratory analysis. *Journal of Consumer Psychology, 5*(3), 209–230. https://doi.org/10.1207/s15327663jcp0503_01.

Jones, R. H. (1997). Marketing the damaged self: The construction of identity in advertisements directed towards people with HIV/AIDS. *Journal of Sociolinguistics, 1*(3), 393–418.

Kaiser Family Foundation (KFF). (2001). *Understanding the effects of direct-to-consumer prescription drug advertising (Document No. 3197)*. Menlo Park, CA: Author.

Kaphingst, K. A., DeJong, W., Rudd, R. E., & Daltroy, L. H. (2004). A content analysis of direct-to-consumer television prescription drug advertisements. *Journal of Health Communication, 9*(6), 515–528.

Kaphingst, K. A., Rudd, R. E., DeJong, W., & Daltroy, L. H. (2005). Comprehension of information in three direct-to-consumer television prescription drug advertisements among adults with limited literacy. *Journal of Health Communication, 10*(7), 609–619. https://doi.org/10.1080/10810730500267647.

Kline, K. N. (2011). Popular media and health images and effects. In T. L. Thompson, R. Parrott, & J. F. Nussbaum (Eds.), *The Routledge handbook of health communication* (pp. 252–267). New York: Routledge.

Kollath-Cattano, C. L., Abad-Vivero, E. N., Thrasher, J. F., Bansal-Travers, M., O'Connor, R. J., Krugman, D. M., et al. (2014). Adult smokers' responses to "corrective statements" regarding tobacco industry deception. *American Journal of Preventive Medicine, 47*(1), 26–36. https://doi.org/10.1016/j.amepre.2014.02.006.

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication, 50*(1), 46–70.

Lee, A. L. (2010). The impact of direct-to-consumer prescription drug advertising research: A diffusion of policy analysis. *Health Communication, 25*(6/7), 619–621. https://doi.org/10.1080/10410236.2010.497466.

Macias, W., Lewis, L. S., & Tae Hyun, B. (2010). The changing face of direct-to-consumer print advertising. *Pharmaceutical Medicine, 24*(3), 165–177.

Mahon, J. E. (2016). The definition of lying and deception. In E. N. Zalta (Ed.), The Stanford encyclopedia of philosophy. Retrieved from http://plato.stanford.edu/archives/spr2016/entries/lying-definition/.

McAllister, M. P. (1992). AIDS, medicalization, and the news media. In T. Edgar, M. A. Fitzpatrick, & V. S. Freimuth (Eds.), *AIDS: A communication perspective* (pp. 195–221). Hillsdale, NJ: Erlbaum.

Menon, A. M., Deshpande, A. D., Perri III, M., & Zinkhan, G. M. (2003). Consumers' attention to the brief summary in print direct-to-consumer advertisements: Perceived usefulness in patient-physician discussions. *Journal of Public Policy & Marketing, 22*(2), 181–191.

Menon, A. M., Deshpande, A., Zinkhan, G. M., & Perri III, M. (2004). A model assessing the effectiveness of direct-to-consumer advertising: Integration of concepts and measures from marketing and healthcare. *International Journal of Advertising, 23*(1), 91–118.

Messaris, P. (1997). *Visual persuasion: The role of images in advertising*. Thousand Oaks, CA: Sage.

Norris, R. L., Bailey, R. L., Bolls, P. D., & Wise, K. R. (2012). Effects of emotional tone and visual complexity on processing health information in prescription drug advertising. *Health Communication, 27*(1), 42–48. https://doi.org/10.1080/10410236.2011.567450.

Paivio, A. (1986). *Mental representations. A dual coding approach.* Oxford, UK: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie, 45*(3), 255–287. https://doi.org/10.1037/h0084295.

Pallegedara, R., & Hancock, J. T. (2014). Drug advertising. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 311–313). Thousand Oaks, CA: Sage.

Polen, H. H., Khanfar, N. M., & Clauson, K. A. (2009). Impact of direct-to-consumer advertising (DTCA) on patient health-related behaviors and issues. *Health Marketing Quarterly, 26*(1), 42–55.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*(2), 157–171.

Reese, S. D. (1984). Visual-verbal redundancy effects on television news learning. *Journal of Broadcasting, 28*(1), 79–87. https://doi.org/10.1080/08838158409386516.

Richards, J. I. (1990). *Deceptive advertising: Behavioral study of a legal concept.* Hillsdale, NJ: Lawrence Erlbaum.

Richards, J. I., & Petty, R. D. (2007). Advertising regulation. In G. J. Tellis & T. Ambler (Eds.), *The Sage handbook of advertising* (pp. 383–397). Thousand Oaks, CA: Sage.

Riggulsford, M. (2013). *Health and medical public relations.* New York, NY: Routledge.

Rodale Inc. (2003). *Consumer reaction to DTC advertising of prescription medicines 1997 to 2002.* Paper presented to the Food and Drug Administration, U.S. Department of Health and Human Services.

Rosenthal, M. B., Berndt, E. R., Donohue, J. M., Epstein, A. M., & Frank, R. G. (2003). Demand effects of recent changes in prescription drug promotion. In O. M. Cutler & A. M. Garber (Eds.), *Frontiers in health policy research* (Vol. 6). Cambridge, MA: MIT Press.

Russell, C. A., Swasy, J. L., Russell, D. W., & Engel, L. (2017). Eye-tracking evidence that happy faces impair verbal message comprehension: The case of health warnings in direct-to-consumer pharmaceutical television commercials. *International Journal of Advertising, 36*(1), 82–106. https://doi.org/10.1080/02650487.2016.1196030.

Scalvini, M. (2010). Glamorizing sick bodies: How commercial advertising has changed the representation of HIV/AIDS. *Social Semiotics, 20*(3), 219–231.

Shah, M. S., Holmes, E. R., & Desselle, S. P. (2003). The use of persuasion in print DTC advertisements of prescription drugs: A content analysis of leading consumer magazines from 1995–2000. *Journal of Pharmaceutical Marketing & Management, 15*(3), 23–43.

Steel, J. (1998). *Truth, lies, and advertising: The art of account planning.* New York: Wiley.

Walma van der Molen, J. H., & Van der Voort, T. H. A. (2000). Children's and adults' recall of television and print news in children's and adult news formats. *Communication Research, 27*(2), 132–160.

Wilkes, M. S., Bell, R. A., & Kravitz, R. L. (2000). Direct-to-consumer prescription drug advertising: Trends, impact, and implications. *Health Affairs, 19,* 110–128.

Wilkins, L., & Coleman, R. (2005). *The moral media: How journalists reason about ethics.* Mahwah, NJ: Erlbaum.

Wright, P. (1979). Concrete action plans in TV messages to increase reading of drug warnings. *Journal of Consumer Research, 6,* 256–269.

Zhong, W., Maradit-Kremers, H., St. Sauver, J. L., Yawn, B. P., Ebbert, J. O., Roger, V. L., … Rocca, W. A. (2013). Age and sex patterns of drug prescribing in a defined American population. *Mayo Clinic Proceedings, 88*(7), 697–707. https://doi.org/10.1016/j.mayocp.2013.04.021.

CHAPTER 44

Visual Deception: From Camo to Cameron

Paul Martin Lester and Marjorie Yambor

The art of pleasing is the art of deception.
Luc de Clapiers, 1715–1747
French author, moralist, and a friend of Voltaire (Wallis, 1928)

For a mere four dollars and 95 cents, a magnet can be purchased from Amazon.com of de Clapiers' quotation that starts this discussion on visual deception. Although his only contribution to the world's literature was an anonymous volume published a year before his untimely death, the link between pleasure and deception as expressed by a largely unknown eighteenth-century writer is the only saying of his that shares refrigerator door space with a child's drawing or a family snapshot. Unlike the other chapters in this handbook, this work considers deceptive displays as disseminated by the mass media through the lens of visual communication. Visual deception is the art and craft of using observable actions and images to deceive viewers. When visual messages and trickery are combined as with camouflage, magic, disguise, optical illusions, trompe l'oeil, motion pictures, and augmented and virtual realities, the experience can be enjoyable. However, when created for nefarious motivations as with advertising, photography, and robotics, the use of deception, at best, can be considered ethically questionable and at worst, dangerous to life and limb.

P. M. Lester (✉)
School of Arts, Technology, and Emerging Communication,
University of Texas at Dallas, Richardson, TX, USA
e-mail: lester@utdallas.edu

M. Yambor
Department of Communication, Dalton State College, Dalton, Georgia
e-mail: myambor@daltonstate.edu

857

Camouflage

Camouflage is a game we all like to play.
Russell Lynes, 1910–1991
Historian, photographer, and managing editor of *Harper's* (Severo, 1991)

Deriving pleasure from camouflage greatly depends on whether you are predator, prey, or casual observer. If a predator, you may become frustrated by the elaborate colorization and pattern effects achieved by such animals as the dead leaf butterfly that has wings that fold to resemble, well, a dead leaf, the Indonesian mimic octopus and the cuttlefish that can change colors and textures to blend with backgrounds, the Malaysian orchid mantis that imitates a white orchid common to the region, the leafy sea dragon that resembles a floating strand of kelp, or American soldiers in Afghanistan who wear uniforms that are meant to merge with the desert environment (Pegg, 2014). Think of former California Governor Arnold Schwarzenegger's difficulty in spotting the alien sport hunter in the 1987 motion picture aptly named *Predator*. In the first movie of the popular series, the principles of biology combined with advanced technologies achieve the ultimate camouflage—invisibility. If you are prey, you should be thankful for the genetic and/or innovative qualities that render you, when desired, difficult to see by others regardless of their motives. Finally, if you are a casual observer, deception through camouflage can produce an "aha" visual experience whenever a foreground element suddenly reveals itself from a background.

Hollywood producers have tried to entertain audiences with the benefits and pitfalls of being completely camouflaged with films that on the whole should have remained invisible to moviegoers. An exception is the first movie of the genre. Based on the novel by H. G. Wells, the 1933 classic *The Invisible Man* starred Claude Rains. Directed by James Whale, who two years earlier introduced audiences to Dr. Frankenstein's monster, it showcased innovative visual effects created by John Fulton and his team. Unfortunately, other films were not as thoughtfully produced and included such duds as *The Invisible Woman* (1940), *The Invisible Boy* (1957), *Mr. Superinvisible* (1970), *Memoirs of an Invisible Man* (1990), and *Hollow Man* (2000). There is also an important plot twist that involves an "invisibility cloak" that first appeared in J. K. Rowling's *Harry Potter and the Philosopher's Stone* (Vincent, 2016). A popular dinner party topic inspired after an episode of Ira Glass' "This American Life" is whether you would want the power to fly or the ability to be invisible ("Superpowers," 2001). Being able to fly while invisible was not an option; however, psychologists have theorized that the choice you make indicates whether you are the type of person that is immersed and involved in the world around you or content to watch as the day passes.

In the real world, camouflage is a combination of minimizing shadows from light sources so that a three-dimensional object has the appearance of being as flat as possible while the figure–ground relationship becomes one.

The Danish gestalt psychologist Edgar Rubin developed a principle of camouflage when he made patterns with little or no separation between the foreground and the background. Understanding and manipulating this trait of visual perception led directly to military applications that merged the colors and shapes of uniforms and equipment with those of surrounding backgrounds in order to hide. Predators adopted camouflage to avoid being prey in order to be better predators.

Magic

> Not only do I lie, I take real pleasure in lying, in the transmission of magic effects.
> *Ricky Jay, 1948–*
> Stage magician, actor, and author (Jay, 2016)

Unlike camouflage which is the art of prey-avoidance, magic is a form of deception in which the quarry *wants* to be the prey. Fortunately, being awed and perhaps inspired by an able prestidigitator is a victimless crime. The reason why magic is a form of amusement for most may be because there is an implied social contract between performer and participant that requires observing without cynicism and deceit. The viewer must also accept the proposition that "magic creates the illusion of an impossible reality" (Tempest, 2014). The level that one readily accepts being fooled and finds joy in a magician's manipulations of the laws of physics is a journey in the understanding of humbleness. Magic, then, is a powerful, personal ego barometer.

Author, actor, and performer Richard Jay Potash, known professionally as Ricky Jay, relies on his imagination, experience, and an extensive knowledge of the lives and practices of famous magicians throughout history to elevate his art to the point that he has been called "the most gifted sleight-of-hand artist alive" (Singer, 1993). Marco Tempest, a Swiss stage magician most known for his tricks that involve the use of augmented reality technology on smartphones, also has a day job as a director's fellow at the prestigious Massachusetts Institute of Technology. He explains the reason magic has been popular for several centuries with a simple maxim, "It's fun to be fooled" (Tempest, 2014). Unlike crowd-pleasing illusionists such as David Copperfield who made a Learjet, an Orient Express dining car, and the Statue of Liberty disappear, personal, close-up magicians such as Jay and Tempest use deception, misdirection, and dexterity to subtly and quietly astound and amuse.

Several Chicago bars and clubs used to be havens for coin and card manipulators, who doubled as bartenders. With names such as Little Bit O' Magic, and Houdini's Pub and Pizza Magic, eager patrons waited for a spot at the bar, ordered a drink, and for the price of a tip, were fooled ("A Modern Take," n.d.). The New York Lounge was a smoky, dark, red velvet haven for sleight-of-hand wonderment in the tradition of Ricky Jay. A smooth-talking

yet friendly bargician laid a towel in front of you and with cups, balls, coins, cards, and agile fingers, made objects vanish despite your best effort to not be deceived.

Marco Tempest updates sleight-of-hand, close-up magic with iPhones, clever apps, a fast-running video camera, and a projector to perform disappearance and alteration magic before large audiences. His TED talks have been viewed online by millions. With a friendly, matter-of-fact patter, Tempest lulls an audience into a state of deception in which one could truly believe that magic actually exists. His tricks prompted Princess Stephanie of Monaco to exclaim, "I don't know what magic will be like in 50 years, but I suspect that it will look a lot like Marco Tempest." For Tempest (2014), "Art is a deception that creates real emotions—a lie that creates a truth. And when you give yourself over to that deception, it becomes magic." Magic is a combination of art and technology designed to serve deception. The end result should be entertainment.

Disguise

In the theater lying is looked upon as an occupational disease.
Tallulah Bankhead, 1902–1968
Stage and screen actress, libertine (Vandenbroucke, 2001)

Theater sets the stage for the celebration of deception. The mechanics of the theatrical performance invite an illusion shared by actors and audiences alike. House lights dim and stage lights rise to direct focus to a central space of elocution, location, and action. The most common conception of disguise, of course, is the costume—which erases the actors and conjures the cast. English Renaissance theatre prohibited women from appearing on stage, so men played the female roles. Shakespeare's works featured many female characters who masqueraded as men (e.g., Julia in *The Two Gentlemen of Verona*, Portia in *The Merchant of Venice*, and Viola in *Twelfth Night*), so there was the "double deception of a boy playing a girl playing a boy" (Garber, 2004, p. 49). With clothing, styling, makeup, masks, and posture, a human of the real world transforms into a character of a fake realm. Complementary language and delivery lace the visual illusion with authenticity.

Although the tangible size of the stage is circumscribed and stationary, the settings within the narratives are wide and fluid, limited only by the playwright's imagination and designated by sound, scenery, lighting, and stage marking. Andrew Lloyd Webber's rock musical *Starlight Express* (a dream deception, since the story takes place inside the sleep-state of a child's mind) originally featured skate tracks throughout the auditorium that extended and elaborated the sensation of setting, allowing human actors to perform as train engines and railroad cars. Additionally, individual acts within plays trick time as events unfold in a distorted dance of collapsed chronology. Movement onstage

offers the semblance of autonomy yet is confined to the narrative set forth in the script, maneuvered via marks and prompts to propel action in the story.

Accommodating the human thirst for and fascination with themes of trickery and tragedy, dramatic classics reveal narratives rich with betrayal and bamboozlement. Iago in Shakespeare's *Othello* asserts his reputation as an honest man, and yet he delights in deceit—even disclosing, "I am not what I am" (*Othello*, 1.1.66). Molière's classic *Tartuffe* tells the tale of an alluring imposter of the same name, cloaked in charisma with pious pretense, who climactically confesses, "Ah, no, don't be deceived by hollow shows; / I'm far, alas, from being what men suppose" (Wilbur, 1965). Playwright David Ives recently adapted a classic French Restoration comedy by Pierre Corneille entitled *The Liar* that features the character Dorante, who confuses two female friends, concocts an elaborate invention to avoid marrying the wrong woman, and consequently "launches the ensemble of characters into a fast-paced comedy full of mischief, misunderstanding, and mistaken identity" ("The Western Stage," 2015, p. 1).

Identity is constructed on the stage of everyday life as well as in the theatrical arena. In communication theory, Kenneth Burke's work with dramatism suggests that life *is* drama (Griffin, Ledbetter, & Sparks, 2015). Messages are choreographed expressions that convey human motivations, with language as the driving device behind acts, scenes, and agents. Burke describes God-terms and Devil-terms as specific, recurrent positive and negative words used to set the tone and establish themes of speech in search of rhetorical redemption (Griffin et al., 2015). This sociopolitical rhetoric results in a visualization of victimage and identifies images of cultural categories of external enemies that keep the mainstream from flourishing. Examples include race (shade of skin), class (conspicuity of consumption), gender (fashion of physique), religion (attire of affiliation), occupation (uniform of achievement), and sexual orientation (portrayal of partners).

As with the theater, life is also a form of drama—a form of disguise. In 1947, New Orleans photographer Clarence John Laughlin, who began his career as a poet, made a picture he titled, "The Masks Grow to Us." The image is a multiple exposure of a doll's face overlaying the face of a model. His caption reveals his intended meaning: "In our society, most of us wear protective masks (psychological ones) of various kinds and for various reasons. This process is indicated in visual, and symbolic, terms here by several exposures on one negative—the disturbing factor being that the mask is like the girl herself, grown harder and more superficial" ("The Masks," 2017). Just as actors learn the ways to communicate their deceptive craft mostly for the good, others with hidden agendas, covert motives, and a need for self-preservation sometimes use deceit as a way of constructing self. For example, media often portray transpersons as needing to "pass" as their chosen gender using outlandish makeup and costumes and dramatizing bizarre behaviors that support preconceived stereotypes. Passing as a concept became

necessary when persons considered as "others" wanted to participate in the benefits afforded those of the dominant culture without detection. Passing may be required because of perceived differences in race, class, religion, gender, or another identity among dominant and non-dominant individuals. In dark-skin cultures, it is called "colorism." For example, Asian persons have learned that if they have lighter skin they will have an appreciable advantage over those who are darker (Mariam, 2017). Other cultural groups have their own forms of passing. In F. Scott Fitzgerald's *The Great Gatsby*, those without economic means try to blend with the wealthy classes. Religious passing has a more sinister history. For example, the motion picture *Europa Europa* (1990) tells the story of Jewish men who attempted surgery to restore their foreskins to pass as Gentiles in order to escape the horrors of the Holocaust. Passing, then, is a thoroughly visual phenomenon, not only for transpersons, but for other groups as well. As such, passing is yet another piece in the deception puzzle. These examples are disguises by the nature of their social construction. Whether on the stage of entertainment or the stage of life, disguise deploys deception via imagination and integration.

Optical Illusions

Illusion is the first of all pleasures.
Voltaire, 1694–1778
Writer, philosopher, historian (Douglas, 1917)

Optical illusions reenvision reality by perverting perception. The eye sees, the brain processes, and the mind interprets objects as different than what they actually are. Scientists categorize optical illusions according to three types: literal, physiological, and cognitive. Literal illusions include collections of images that form a separate holistic object, so the eyes can shift focus among the parts and the whole (e.g., Charles Allan Gilbert's *All Is Vanity*). Feeding from stimuli such as size, color, and position, physiological illusions engage a sensory overload that triggers an afterimage that does not exist (e.g., The Hermann grid by Ludimar Hermann). Cognitive illusions challenge assumptions and predispositions about the nature of and rules for objects in the physical world (e.g., Penrose's impossible staircase) (Lester, 2011).

All illusions arouse some sort of psychological sensation that invokes deception. Hermann Rorschach created the first systematic approach to explore personality dynamics by assessing reactions to his ink drawings. With its infamous inkblots, the Rorschach-test imagery has played a role in many media moments, such as Andy Warhol's art (*Rorschach*), print advertising (AXE, Saab, Mastercard), and Hollywood films (*The Mothman Prophecies*, *Superbad*, *Watchmen*). The pop culture portrayal, however, misrepresents how the actual test was intended to perform and what data it revealed. The media illusion of inkblot interpretation deceives popular perception and reduces a psychological tool to a cultural meme.

Perhaps no optical illusion in recent history has prompted such social media buzz and chaotic cultural confusion as "The dress." Black and blue or white and gold? That was the question that a stirred a deluge of debate about the deceptive nature of a photograph of a dress posted to Tumblr on February 26, 2015. "Dressgate" went viral and commanded international attention among friends, families, celebrities, news agencies, and even neuroscientists. As Stephen Macknik (2015) noted, the scientific community became immediately interested since "The Dress is the very first color illusion in which different people from the neurotypical population—those people with no known differences between their brains—experience the same colored surfaces as drastically different colors." According to some initial studies, variations in color cozenage depended on variables such as age, gender, light distribution, and perceived ambiguities in luminance levels (Macknik, 2015).

The attraction of optical illusion is not limited to the pop science realm. It also satisfies simple visual amusement. The back page of *People* magazine forever featured fabricated photos of the envisaged offspring of two random stars. Celebrity face video mashups have appeared on YouTube (FaceMashups) and on Conan O'Brien's late-night talk show in a segment called "If They Melded," which Conan describes accordingly: "This is where we take videos of famous people and waste everybody's time by combining them. There's no reason to do it" (O'Brien, 2014). Rendering traditional cross-fade techniques obsolete, software such as After Effects, Morpheus, and MorphThing follow the trend of user-generated content and invites everyone to get in on the fun. The art of optical illusion bends the brain in a dance of delightful deception.

Trompe L'Oeil

What you understand no longer matters.
Jane Hirshfield, 1953–
American poet, essayist, translator (1988)

Trompe L'oeil (French for "deceive the eye") describes an art technique and tradition that renders an illusion of three-dimensional space. With their "refusal to respect the limits of the frame," the best versions will trick viewers in such a way that they cannot distinguish that which is painted (or photographed) from that which is real (Ndalianis, 2000). Artists such as Salvador Dali, Rene Magritte, Henri Cadiou, and Banksy have composed in the trompe l'oeil style. Images include semblances of height, depth, curve, texture, stone, carving, columns, pillars, arches, objects, implements, landscapes—really, anything imaginable. Paint strokes on furniture imitate random articles left lying around. Sidewalk-chalk sketches draw the eye to the foot and shift the step. Murals adorn interiors and exteriors in twists of context. Architectural trompe l'oeil features perspective painting and opens up space (typically walls or ceilings) by simulating actual windows, doors,

balconies, and skylights. This style proved popular in baroque cathedrals to add impressive visual interest to otherwise simple structures, creating an illusion of opulence. As van de Port (2012) describes:

> Expressing infinity, these ceilings underscore the gap between God and humankind. Yet, simultaneously, they speak to the desire to contact that unreachable God. Indeed, looking up to these ceilings, registering the slight dizziness they induce, the possibility to be connected to the infinite becomes sense-able. (p. 874)

Connection is also critical for audience engagement and enjoyment in stage, film, and television productions. A large part of that begins with believable sets that transport the mind into the narrative space; trompe l'oeil satisfies the standards for three-dimensional design to achieve this effect. Scenes with this sensation can be seen in films such as *Eternal Sunshine of the Spotless Mind*, *Who Framed Roger Rabbit*, and *Singin' in the Rain*. In some instances, the pop culture texts themselves feature trompe l'oeil within the storyline: Wile E. Coyote paints tunnels onto rock walls (which accommodate Road Runner but foil Coyote) in *Looney Tunes* cartoons. The television series "Dallas" and "Westworld" dedicate entire episodes to the trompe l'oeil theme (see "Trompe L'Oeil," 1986, 2016).

Reaching beyond renowned public spaces and popular culture, trompe l'oeil allows society's well-to-do sector to access the appearance of affluence in their homes and lives by recreating rare objects available only to the wealthiest. Items that one owns have always signified social status. As Puşcaşiu (2015) points out, "materials are carriers of certain information, and faking them speaks volumes about the socio-economical context in which the objects were created and about their respective owners" (p. 114). One automatically thinks of trompe l'oeil-gone-tangible: the fake watches, knock-off handbags, and counterfeit clothing sold by street vendors in major cities all over the world that are "Made to deceive and look like something they are not, the spell [is] mostly broken when the object [is] touched, or even upon closer visual inspection, which must mean that the trick was not meant to be long-standing, the lie was not meant to be permanent" (Puşcaşiu, 2015). However, whether it is an art object forged on a wall, a film scene projected onto a screen, or a designer handbag bought on the street, the initial impact of the image deceives by inspiring the expectation of the original—especially powerful in the present-day at-a-glance essence that all too often refuses to delve deeply to locate authenticity and believability.

Advertising

If you live in a world full of politicians and advertising, there's obviously a lot of deception.
Kenneth Koch, 1925–2002
Poet, author, educator ("Kenneth Koch," 2017)

Advertising has been called the world's second oldest profession, and like its cousin, public relations rely primarily on images to aid in its mission to persuade. Unfortunately, the profession is criticized when it also involves deception—when honest persuasion shifts to outright propaganda. The importance of visual messages in the persuasive process, for good or not, was expressed in 1922 by the American journalist and media critic Walter Lippmann. He stressed the need for images to change a person's attitude. In *Public Opinion* Lippmann (1922, p. 162) wrote, "Pictures have always been the surest way of conveying an idea and next in order, words that call up pictures in memory."

Perhaps inspired by his endorsement and continuing the successful propaganda poster tradition from World War I, government officials, advertisers, and graphic designers embraced visual deception in the form of colorful and visually eye-catching political posters. Examples include James Montgomery Flagg's poster of a finger-pointing Uncle Sam, J. Howard Miller's "We Can Do It!" that shows a woman factory worker, and another powerful picture, Norman Rockwell's "Rosie the Riveter" (Lester, 2011).

However, in terms of visual deception during this era, no one did it better than the Nazi regime. In 1933, the Third Reich of Nazi Germany established the Ministry of Propaganda with Joseph Goebbels as its head responsible for film and poster production. Typical was a piece labeled "LIBERATORS" that was a complicated collection of visual symbols used in misleading ways. A Ku Klux Klan hooded robotic giant meant to be the US is in the act of destruction. With two muscular arms, one holding an LP record and the other a money bag, and two additional arms with one holding a machine gun and the other a grenade, the poster alludes to America's presumed racism, Jewish sympathies, and obsession with beauty, consumerism, and entertainment. Prone citizens near a town's traditional water fountain await the onslaught of American music, superficial beauty, racism, Jewish interests, and violence. In this symbolic-laden World War II propaganda poster, American culture is feared if Germany loses the war (Lester, 2011).

Although certainly not on par with German propaganda, advertising is nevertheless occasionally critiqued for misleading campaigns that can have dire consequences for unsuspecting consumers. Will Heilpern (2016) writing for *Business Insider* magazine noted "18 false advertising scandals" that must lead consumers to the conclusion that the profession is infused with an ethos of deception in its story-telling and image selection choices. The actress Jamie Lee Curtis sincerely promised in commercials for Dannon's Activia yogurt brand that the product would boost immune systems. Red Bull showed a picture of a model in an astronaut's costume sporting wings that implied that the drink improved concentration and intelligence. Olay used the 1960s model Twiggy to show that its Definity eye cream could make you look younger if you employed Photoshop technicians to touch up your face for the close-up portrait.

Glenn Ruppel and Ruth Reiss (2009) of *ABC News* wrote a "20/20" investigation of photographic deceptions used by the travel industry to entice

travelers to stay at specific hotels or resorts. Photographs on websites were manipulated so that hotels looked closer to popular landmarks than they were, areas full of guests look deserted, quiet, and private, fitness centers were much more spacious than they appeared, and rooms were more modern and nicer than when you checked in. Most egregious was a picture of a quaint hostelry in Poland under a clear, blue sky. Upon arrival, the place is located on a busy street with a huge power plant's cooling tower next door. The unsightly detail was easily removed through Photoshop. Bill Hilson, professor at the Pratt Institute, explained to a reporter, "I wish I could tell you it was incredibly sophisticated and difficult to do, but today it is not." Hilson then noted a more important issue, "I know for a fact that virtually every single image used in advertising and marketing today has been adjusted to some degree" (Ruppell & Reiss, 2009). Substitute deceptive for adjusted.

PHOTOGRAPHY

Photography has always been capable of manipulation.
Joel Sternfeld, 1944–
Fine art photographer and educator (Keats, 2012)

The photographic profession has been called a good choice for artistically minded and technically unchallenged individuals who, nevertheless, are afflicted with short attention spans. The online Shaw Academy ("Top 15 Genres," 2016) acknowledges this condition with its pitch to students to learn the "Top 15 Genres of Photography That You Need to Know." The 15 featured are: aerial, architectural, candid, documentary, fashion, food, landscape, night-long exposure, photojournalism, conceptual/fine art, portraiture, sport, street, war, and wildlife. Never mind that some of the categories overlap. Once you learn the basics of subject selection, composition, the decisive moment, camera and lens choices, manual controls, lighting techniques, and so on, if you get bored with one type of image making, you can easily move to another. Interestingly, in all of the genres mentioned, deception has either been a critical ethical concern in the past or an accepted practice in the present.

Transformed by the use of drones, aerial photography as realized by Turkish art photographer Aydın Büyüktaş uses manipulation software to create deceptively acceptable "dream-like multidimensional roller coasters" of locations in his native country and in the US for his *Flatland* projects (Koblyakova, 2017). The web offers many examples of how-to sites and YouTube videos to explain the procedures for making food appetizing in pictures (Zhang, 2016). Most wildlife pictures are faked. Salina Cheng and Hannah Yi (2017) write of a Minnesota game farm that invites photographers for up to $1000 each to take pictures "in the wild" of their pets that include "a black bear, five cougars, 20 wolves, 46 foxes, and a 2,000-pound bison."

Deception is a crucial element in many genres of photography with viewers willing to suspend their disbelief in order to be visually entertained. Most readers, however, object to trickery in documentary, photojournalism, and war photography. Dorothea Lange used her formal portrait skills learned as a photographer in her San Francisco studio to stage-manage Florence Thompson and four of her children in 1936. Lange created one of the most enduring images from America's Great Depression titled, "The Migrant Mother" (Lester, 2011). Most photojournalists today would be fired for manipulating the family to such an extent, particularly for a news event. Two recent photojournalists experienced career setbacks for their digitally deceptive news pictures. In 2003, Brian Walski, a staff photojournalist for the *Los Angeles Times* created a photograph composite while he covered the war in Iraq that combined two different images into one. In one picture, a British soldier gestures toward a group of men. In the other photograph, a man walks while holding a child. The manipulated picture was printed on the front page. After the deception was discovered, Walski was fired. Photojournalist Souvid Datta in 2017 copied part of a picture from Mary Ellen Mark's classic 1981 documentary *Falkland Road* and pasted it in the background of one of his own images (Lester, 2018).

Lange's deception is excused because of an ethical concept known as time forgiveness. Cultural values and professional practices change over decades, especially when a photograph is as treasured as the Thompson portrait. Walski and Datta, living during the present time, are not so readily excused for their unethical behavior. After the deception was discovered, Walski was fired from his newspaper position while Datta's reputation was shattered. He took down his website and discontinued his Facebook and Twitter accounts.

ROBOTS

Human beings have dreams. Even dogs have dreams, but not you; you are just a machine. An imitation of life.
Del Spooner AKA *Will Smith*
Detective and warrior (Smith, 2004)

Robots mimic a human's mind and body by their ability to complete tasks with accuracy, agility, and efficiency. While those in research labs and manufacturing plants focus on functionality, those that possess personality have maintained memorable roles in pop culture consciousness. Through oral legends, written records, and actual objects, the construction of humanoid replacements has a long history fooling an unsuspecting public in elaborate deceptions. There were myths that dragon teeth could turn into soldiers, lifelike statues that protected palaces, mechanical pigeons powered by steam, and mechanical musicians that floated on water from Greek, Arab, Italian, and Japanese creatives (Gera, 2003; Godwin, 1876; Sharkey, 2007; Hornyak, 2006). Leonardo da Vinci in one of his notebooks described a mechanical

knight ("A Brief History," 2005). Japanese inventors created elaborate automated toys. Arguably the most famous contraption was the "Turk," an elaborate hoax created by the Hungarian mathematician Wolfgang von Kempelen in 1769. The chess-playing automaton dressed in Turkish garb and sat behind a large cabinet that held a chessboard. Kempelen performed with his device throughout the world and won games of chess against the likes of Napoleon Bonaparte and Benjamin Franklin. In reality, a small man hid inside the cabinet pulling levers to simulate the robot's moves (Standage, 2002).

In more recent times, examples abound. Rosie keeps house in "The Jetsons"; Optimus Prime leads the *Transformers* franchise, and R2-D2 and C-3PO enhance the *Star Wars* sensibility. Designers frequently fashion robots with an anthropoid anatomy, automatically establishing an expectation of human interaction and opening an avenue of deception since machines can never be human. Or can they? In the world of technology, magnetized fluids used by NASA lubricate robot parts to allow more natural movement, and artificial multifilament muscles developed by Japanese researchers mimic the muscular system. Synthetic humans (SynAtomy, SynTissue, SynDavers) substitute for humans in medical research, training, and simulation (SynDaver Labs, 2017). Researchers in Switzerland discovered that "robots equipped with artificial neural networks and programmed to find 'food' eventually learned to conceal their visual signals from other robots to keep the food for themselves" (Grifantini, 2009). The robots "communicated" with one another by randomly flashing a blue light, but as they evolved via computer code that combined the most successful neural networks sprinkled with random biological mutations (engineered survival of the fittest), they rarely flashed their lights when they were near food so they could earn more points and amass the food. Agents of artificial intelligence learned how to dupe their peers and veil their visual cues of success to survive and thrive, as if they had heeded the advice of Frank Lucas (played by Denzel Washington) when he admonishes his brother Huey for donning a flashy suit in *American Gangster*: "That's a clown suit. That's a costume, with a big sign on it that says 'Arrest me.' You understand? You're too loud; you're making too much noise. Look at me: the loudest one in the room is the weakest one in the room" (Washington, 2007).

No one wants to hold the weaker hand in the game of deception. Isaac Asimov's "Three Laws of Robotics" lays the groundwork for amicable human-robot interactions. Generally, the laws of this science fiction universe assert that robots "may not harm humanity, or, by inaction, allow humanity to come to harm" (Luokkala, 2014). Contemporary machine ethicists and artificial intelligence safety researchers consider how to develop artificial superintelligence (ASI) in ways that will avoid a societal breakdown as explored in *Robopocalypse* by Daniel H. Wilson (2011). HBO's "Westworld" probes these prospects in its theme-park settings where the critical characteristics of the robots extend beyond the anthropomorphic bodies: "they display

emotion including extreme pain, they see and recognise each other's suffering, they bleed and even die." The visual of blood—"'robot blood,' a classic science fiction trope designed to confound mechanical expectations"—invites a visceral response from the viewer and seals the sophism of the robot-as-sentient-being (Riley, 2016). While the uncanny valley hypothesis suggests that human-robot verisimilitude is too close for comfort, researchers such as Adriana Hamacher reveal that people empathize with and relate to robots that display emotional expressiveness (Velocci, 2016). Despite the deception inherent in the corporeal-mechanical interaction, in both technology and entertainment, robots evoke fascination with the deceptions of the human-robot relation.

Augmented and Virtual Realities

Reality is merely an illusion, albeit a very persistent one.
Albert Einstein, 1879–1955
Physicist and philosopher ("Albert Einstein," 2017)

Writing in *Telepolis* magazine (a German internet magazine), Goedart Palm (2005) introduces the reader to the nineteenth-century German satirist and futurist, Alexander Moszkowski. As a friend of Einstein, Moszkowski was one of the first authors to help publicize the physicist's theory of relativity. He was also known as a humorist who used his wit to parody social customs and employed his understanding of scientific principles to predict technological innovations that are commonly used today. In his 1922 science fiction novel *The Isles of Wisdom*, he described a series of islands in which various forms of utopian and dystopian worlds prosper or decline. In the book, he predicted today's information-driven society, the ubiquitous nature of smartphones, and virtual reality (VR), a name invented by new media guru Jaron Lanier.

In 1968, the American computer scientist Ivan Sutherland, known as the "founder of computer graphics" while a professor at Harvard and with his student Bob Sproull, brought Moszkowski's vision to life. They invented one of the first head-mounted virtual reality displays. It was composed of a simple, stereoscopic image that required a "mechanical arm suspended from the ceiling of the lab" (Rheingold, 1992). By 1992 VR became a cultural meme after Hollywood offered its unsettling versions of mind-altering displays as seen in *The Lawnmower Man*, *The Matrix*, *Virtuosity*, and *Strange Days*. Presently, news and entertainment companies as well as academic centers offer less fretful presentations. Media entities from *ABC News* to *The New York Times*, games from HTC Vive, Oculus Rift, and PlayStation VR, as well as educational institutions such as the Newhouse School at Syracuse University, Stanford University's Virtual Human Interaction Lab, and the School of Cinematic Arts at the University of Southern California (USC) have created critically acclaimed VR motion pictures (Lester, 2018).

Closely related to VR is augmented (AR) or mixed reality (MR). Google Glass was the highly anticipated AR eyewear that displayed digital interfaces that blended with a user's location. The internet and other features were accessed through finger and voice commands. However, with concerns about privacy (facial recognition software could identify strangers on the street), health (eye strain and distracted walking and driving), and ethics (interviewees not knowing their words and actions were recorded by Glass journalists), the prototype was discontinued in 2015. In the meantime, Microsoft has introduced a prototype currently named HoloLens. It's a pair of mixed reality smartglasses within a holographic platform. Through simple hand gestures, users can manipulate and work on any file as well as watch movies and play games (Lester, 2018).

It is quite possible that AR will eventually overtake VR, but in the meantime, virtual reality dominates news stories and the public's imagination. The reason for the recent interest is the combination of deception and entertainment. Not surprisingly, Ivan Sutherland (1965) named his version "the ultimate display." The term is an apt description that brings to mind a statement about innovation from Arthur C. Clarke (2000), "Any sufficiently advanced technology is indistinguishable from magic." Another way to express Clarke's passage might be, "Deception in the form of an immersive computer simulation is indistinguishable from reality." Virtual reality is the ultimate deception engine. What can be more illusory than a technology in which there is no discernable difference between the cybernetic and the existent?

Motion Pictures

Your imagination can create a reality.
James Cameron, 1954–
Filmmaker, screenwriter, deep-sea explorer (Duarte, 2010)

The lights dim. The screen flares. The movie begins: cue the imagination. Motion pictures deliver a decadent depiction of deception, inviting audiences to subdue logic, suspend disbelief, and surrender emotion to embrace the excitement and escapism of the entertainment experience. Through the larger-than-life lens that cleverly combines discreet detachment with an immersive environment, viewers encounter new personalities and explore other worlds. Jon James Miller (2016) recounts the perspective presented to him by his professor in film school: "The burden…is first on the writer to create a compelling enough universe that the audience would willingly forget their own reality."

Poet Samuel Taylor Coleridge coined the term "suspension of disbelief" in 1817. Researchers still know relatively little about the mental mechanics that allow the human brain to accept the action onscreen as real yet not mobilize the motor skills to participate in the scenes. As the limbic system triggers emotional responses, though, Norman H. Holland (Mueller, 2014) notes

that people "are able to believe in the supernatural occurrences in Coleridge's *Ancient Mariner*, the inhuman strength and speed of Superman, or the harrowing journey of a Hobbit in his quest to destroy an evil ring." This affective reaction shifts visual focus. Research reveals that high levels of storyline suspense reduce activity in peripheral visual processing regions and increase activity in central visual processing areas, confirming "dynamic spatial tuning of attention…due to narrative context" (Bezdek et al., 2015).

In 1993, *Jurassic Park* astounded moviegoers and solidified the concept of the suspension of disbelief with computer-generated dinosaurs on the screen so believable that many were deceived and thought they had to live somewhere in this world. Another stand-out example came in 2001 with the hit *Monsters, Inc.* The movie featured several innovations in computer animation technology, mostly related to how hair and clothing were rendered. Animators painstakingly re-created about three million strands of the character Sulley's fur coat. The coding took two years to write to achieve a natural and lifelike appearance. Director John Lasseter wanted theatergoers to say, "Oh, we know it's not real, but it sure does look real" (Lester, 2011).

Nevertheless, expensive animated productions have been rejected by moviegoers who considered the screen characters eerie and off-putting when they reproduced humans too closely. A hypothesis related to life-like robots and animated characters, "the uncanny valley," states that if human imitations seem too exact, they will cause a feeling of revulsion. The concept came from the work by the Japanese roboticist Masahiro Mori in 1970 that involved human responses to non-human creations (Reichardt, 1978). For example, of the 2001 motion picture *Final Fantasy: The Spirits Within*, Peter Travers of *Rolling Stone* wrote, "At first it's fun to watch the characters …. But then you notice a coldness in the eyes, a mechanical quality in the movements" (Travers, 2001). Castle Rock Entertainment lost a fortune on the 2004 *The Polar Express*, which one reviewer described as having characters that were "creepy" and "dead-eyed" and called the *Express* "a zombie train" (Anderson, 2004). About Steven Spielberg's *The Adventures of Tintin: The Secret Life of the Unicorn*, Daniel Snyder (2011) of *The Atlantic* noted, "While all the characters sport some kind of cartoonish features—especially their ears and noses—their photorealistic eyes are somehow blank…. In bringing them to life, Spielberg has made the characters dead." It seems the eyes are the windows to the uncanny valley's floor.

The genius of James Cameron (known for blockbusters such as the *Terminator* films, *Aliens*, and *Titanic*) is that his efforts never journey into the "uncanny valley," usually associated with robots, because they always stress a strong narrative. His 2009 release of the innovative sci-fi epic *Avatar* transcended the boundaries of traditional film technology. The combination of live-action sequences and digitally captured performances "was supposed to be a sort of digital sleight of hand—a human character inhabiting an alien body [to] blend into an alien world, played by a human actor inhabiting a

digital body in a digital world." The majority of Avatar is computer-generated and features elements such as "six-legged hammerhead thanators, armored direhorses, pterodactyl-like banshees, hundreds of trees and plants, floating mountains and incredible landscapes, all created from scratch," based on Cameron's imagination that challenged production parameters with technology such as motion capture suits, scale models, swing cameras, digital animation, and performance-capture sets. To counteract the uncanny valley, tiny cameras mounted on helmets tracked the most minute muscle movements which then circulated in a sea of computer code until the algorithms camouflaged the facade enough to satisfy Cameron's suspension of disbelief (Thompson, 2010). Perhaps motion pictures like *Avatar* invite—even urge—audiences to consider their own liminal lives, increasingly experienced via virtual venues that feel real but are actually addictive, elusive, and deceptive simulations of being-in-the-world.

CONCLUSION

> Nothing is so boring as having to keep up a deception.
> *E. V. Lucas, 1868–1938*
> Author and publisher ("Edward Verrall Lucas," 2017)

As with other examples in this chapter and book, deception, when employed to engage users for positive purposes, can be easily defended. Deceptive practices found through camouflage, magic, disguise, optical illusions, trompe l'oeil, augmented and virtual realities, and motion pictures are usually easily defended and often considered praiseworthy. However, deception in advertising and journalism, photography, and robotics is often criticized and blameworthy. Being deceptive for personal, economic, and political reasons invokes a hedonistic, self-centered philosophy. Producers should strive for a utilitarian, a golden mean, or a veil of ignorance approach that, respectfully, refers to increasing user knowledge, aims to avoid extremes, or helps increase compassion and empathy. If so, even the most deceptive visual messages can be defended and praised.

REFERENCES

A Brief History of Robotics. (2005). Accessed June 15, 2017 from http://goo.gl/wnYVMM.

A Modern Take on a Lost Chicago Tradition. (n.d.). Accessed June 7, 2017 from http://goo.gl/cxDgGW.

Albert Einstein. (2017). *Biography*. Accessed June 8, 2017 from http://goo.gl/O16X6s.

Anderson, J. (2004, November 9). 'Polar Express' derails in zombie land. *Newsday*. Accessed June 22, 2017 from http://goo.gl/BygmD9.

Bezdek, M. A., Gerrig, R. J., Wenzel, W. G., Shin, J., Pirog Revill, K., & Schumacher, E. H. (2015). Neural evidence that suspense narrows attentional focus. *Neuroscience, 303,* 338–345. https://doi.org/10.1016/j.neuroscience.2015.06.055.

Cheng, S., & Yi, H. (2017, April 30). Love wildlife photos? There's a good chance they weren't shot in the wild. *Quartz.* Accessed June 8, 2017 from http://goo.gl/Pb8Egy.

Clarke, A. C. (2000). *Profiles of the future: An inquiry in the limits of the possible.* London: The Orion Publishing Group.

Douglas, C. N. (Ed.). (1917). *Forty thousand quotations: Prose and poetical.* London: George G. Harrap.

Duarte, N. (2010). *Resonate: Present visual stories that transform audiences.* Hoboken: Wiley.

Edward Verrall Lucas. (2017). *Goodreads.* Accessed June 8, 2017 from http://goo.gl/4V1J4n.

Garber, M. (2004). *Shakespeare after all* (p. 49). NY: Anchor Books.

Gera, D. L. (2003). *Ancient Greek ideas on speech, language, and civilization.* Oxford: Oxford University Press.

Godwin, W. (1876). *Lives of the necromancers.* London: Chatto and Windus.

Grifantini, K. (2009, August 18). Robots 'evolve' the ability to deceive. *MIT Technology Review.* Accessed March 14, 2018 from http://goo.gl/unFAKz.

Griffin, E., Ledbetter, A., & Sparks, G. (2015). Dramatism of Kenneth Burke. In E. Griffin, A. Ledbetter, & G. Sparks (Eds.), *A first look at communication theory* (pp. 293–302). New York, NY: McGraw-Hill.

Heilpern, W. (2016, March 31). 18 false advertising scandals that cost some brands millions. *Business Insider.* Accessed June 8, 2017 from http://goo.gl/UmHHIJ.

Hirshfield, J. (1988). *Of gravity & angels.* Middletown, CT: Wesleyan University Press.

Hornyak, T. N. (2006). *Loving the machine: The art and science of Japanese robots.* New York: Kodansha International.

Jay, R. (2016). From the world wide website of Ricky Jay. Accessed June 7, 2017 from http://goo.gl/uBZAd7.

Keats, J. (2012, September 6). Do not trust this Joel Sternfeld photograph. *Forbes.* Accessed June 8, 2017 from http://goo.gl/sFgRAc.

Kenneth Koch. (2017). *Poetry Foundation.* Accessed June 8, 2017 from http://goo.gl/Jh5dWv.

Koblyakova, M. (2017, March 9). Mind-bending drone photo manipulations turn America into a roller coaster. *Creators.* Accessed June 8, 2017 from http://goo.gl/rklSm4.

Lester, P. M. (2011). *Visual communication images with messages.* Boston: Cengage Learning.

Lester, P. M. (2018). *Visual ethics: A guide for photographers, journalists, and filmmakers.* New York: Routledge.

Lippmann, W. (1922). *Public opinion.* New York, NY: Harcourt, Brace, and Company.

Luokkala, B. B. (2014). *Exploring science through science fiction.* New York, NY: Springer.

Macknik, S. L. (2015, May 16). The current biology of the dress. *Scientific American.* Accessed June 22, 2017 from http://goo.gl/9t9p87.

Mariam, S. (2017, May 17). Daring to be dark: Fighting against colorism in South Asian cultures. *Huffpost*. Accessed March 14, 2018 from http://goo.gl/qkxZXG.

Miller, J. J. (2016, May 3). Suspension of disbelief: Making a Hollywood dream a reality. *Script: Division of the writers store*. Accessed June 22, 2017 from http://goo.gl/GdQqaz.

Mueller, M. (2014, January 1). What brain activity can explain suspension of disbelief? *Scientific American*. Accessed June 22, 2017 from http://goo.gl/ZspmVW.

Ndalianis, A. (2000, April). Baroque perceptual regimes. *Senses of cinema*. Accessed June 22, 2017 from http://goo.gl/WURDmy.

O'Brien, C. (Host). (2014, January 28). *Conan* [Television series episode]. Burbank, CA: Warner Bros.

Palm, G. (2005, December 26). Alexander Moszkowski. *Telepolis*. Accessed June 7, 2017 from http://goo.gl/C1L0D3.

Pegg, D. (2014, July 25). 25 incredible camouflaged animals. *List25*. Accessed June 7, 2017 from http://goo.gl/E2cP6V.

Pușcașiu, V. (2015). Design, It's not what it looks like! *International Review of Social Research, 5*(2), 113–119. https://doi.org/10.1515/irsr-2015-0010.

Reichardt, J. (1978). *Robots: Fact, fiction, and prediction*. London: Penguin Books.

Rheingold, H. (1992). *Virtual reality*. New York: Simon & Schuster.

Riley, T. (2016, September 15). 3 theories explaining why robots bleed. *Inverse*. Accessed June 22, 2017 from http://goo.gl/4DJo43.

Ruppel, G., & Reiss, R. (2009, December 11). Beware: Slick web sites use trick photography. *ABC News*. Accessed June 8, 2017 from http://goo.gl/7PMJ0y.

Severo, R. (1991, September 16). Russell Lynes, 80, an editor and arbiter of taste. *The New York Times*.

Sharkey, N. (2007). *A 13th Century programmable robot*. Accessed June 15, 2017 from http://goo.gl/dvZist.

Singer, M. (1993, April 5). Secrets of the Magus. *The New Yorker*. Accessed June 8, 2017 from http://goo.gl/S5sWkd.

Smith, W. (Actor). (2004). *I, Robot* [Motion picture]. USA: 20th Century Fox.

Snyder, D. (2011, December 26). 'Tintin' and the curious case of the dead eyes. *The Atlantic*. Accessed June 22, 2017 from http://goo.gl/UHABQR.

Standage, T. (2002). *The Turk: The life and times of the famous 19th Century chess-playing machine*. London: Walker Books.

Superpowers. (2001, February 23). *This American life*. Accessed June 8, 2017 from http://goo.gl/iwmqxw.

Sutherland, I. (1965). The ultimate display. *Proceedings of IFIP Congress*, 506–508.

SynDaver Labs. (2017). Accessed June 22, 2017 from http://goo.gl/zjzv1E.

Tempest, M. (2014). And for my next trick, a robot. *TED*. Accessed June 7, 2017 from http://goo.gl/eeCerN.

The Masks Grow to Us. (2017). *The Met*. Accessed June 13, 2017 from http://goo.gl/8Dl10c.

The Western Stage. (2015). *The liar: A classic French farce about deception and distortion with a fresh, contemporary spirit* [Press release]. Accessed June 22, 2017 from https://westernstage.com/wp-content/uploads/2011/07/TheWesternStage_TheLiar_pressrelease.pdf.

Thompson, A. (January 1, 2010). How James Cameron's innovative new 3D tech created *Avatar. Popular Mechanics*. Accessed June 22, 2017 from http://goo.gl/jJQof5.

Top 15 genres of photography that you need to know. (2016, February 19). *Shaw Academy*. Accessed June 8, 2017 from http://goo.gl/Dn3xop.

Travers, P. (2001, July 6). Final fantasy. *Rolling Stone*. Accessed June 22, 2017 from http://goo.gl/jPmuWK.

Trompe L'Oeil. (1986, October 24). *Dallas*. Accessed March 14, 2018 from http://goo.gl/23GmGD.

Trompe L'Oeil. (2016, November 13). *Westworld*. Accessed March 14, 2018 from http://goo.gl/LS2Nrt.

van de Port, M. (2012). Genuinely made up: Camp, baroque, and other denaturalizing aesthetics in the cultural production of the real. *Journal the Royal Anthropological Institute, 18*(4), 864–883. https://doi.org/10.1111/j.1467-9655.2012.01796.x.

Vandenbroucke, R. (Ed.). (2001). *The theatre quotation book: A treasury of insights and insults*. New York, NY: Limelight.

Velocci, C. (2016, August 23). Apparently people trusted this robot. *Gizmodo*. Accessed June 22, 2017 from http://goo.gl/ye6Wbc.

Vincent, A. (2016, January 14). JK Rowling explains how Harry Potter got his invisibility cloak. *The Telegraph*. Accessed June 8, 2017 from http://goo.gl/p8jMhr.

Wallis, M. (1928). *Luc de Clapiers, Marquis de Vauvenargues*. Cambridge: Cambridge University Press.

Washington, D. (Actor). (2007). *American Gangster* [Motion picture]. USA: Universal.

Wilbur, R. (1965). *Molière: The Misanthrope and Tartuffe*. New York, NY: Harcourt Brace Jovanovich.

Wilson, D. H. (2011). *Robopocalypse*. NY: Doubleday.

Zhang, M. (2016, March 24). These photos show the secret Tricks of food photography. *PetaPixel*. Accessed June 8, 2017 from http://goo.gl/78RWxq.

Portrayals of Romantic Deception to the Masses: An Analysis of Classic and Contemporary Arts, Modern Technologies, and Empirical Literature

Maria Espinola, Neslihan James-Kangal, Mar Gámez García, Natalie Odisho and Lening Alexis Olivera-Figueroa

When one is in love, one always begins by deceiving one's self, and one always ends by deceiving others. That is what the world calls a romance.

Oscar Wilde

M. Espinola (✉)
Department of Psychiatry and Behavioral Neuroscience,
University of Cincinnati College of Medicine, Cincinnati, OH, USA
e-mail: maria.espinola@uc.edu

N. James-Kangal
Department of Psychology, University of Cincinnati, Cincinnati, OH, USA
e-mail: jamesknn@mail.uc.edu

M. G. García
Department of Romance Languages and Literatures,
University of Cincinnati, Cincinnati, OH, USA
e-mail: gamezgmr@mail.uc.edu

N. Odisho
Florida College of Integrative Medicine, Orlando, FL, USA

L. A. Olivera-Figueroa
Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA

Romantic deception broadly describes many forms of dishonesty that may occur between relationship partners, including lying about their feelings, concealing information from each other, or misrepresenting themselves to gain someone's romantic interests (Peterson, 1996). Generally, romantic deception is seen as a violation of relationship standards, yet some degree of deception exists in most relationships. As with most universal experiences, romantic deception has long been the inspiration for creative works, including literary classics and paintings that date back several centuries, as well as modern television, film, and literature. The dynamics of romantic deception have further evolved with recent developments in technology, such that modern communication outlets increase the opportunities for romantic deception. In this chapter, we discuss both classic and contemporary portrayals of romantic deception across literature, arts, and communication outlets and present an empirically informed analysis of romantic deception based on psychological research.

CLASSIC PORTRAYALS OF ROMANTIC DECEPTION

The classic portrayals of romantic deception can be divided into two main categories: (1) interpersonal deception (e.g., misrepresenting oneself to a romantic interest by concealing information or lying about romantic feelings) and (2) self-deception (e.g., deceiving oneself about the identity of a loved one, the nature of the relationship, or a significant other's romantic feelings).

Classic Literary Portrayals

The most popular classic portrayal in the first category of romantic deception is the story of Don Juan. Since the seventeenth century, the Spanish-born character has epitomized the idea of the male seducer, not only in literature, but also through popular use of his name as a synonym for "womanizer." The character of Don Juan is so popular that it is believed that there exist more than a thousand literary versions of him, without taking into account those of opera, ballet, and cinema (Becerra Suarez, 1997). The most well-known literary versions of the Don Juan character are those of the playwrights by José Zorrilla (1884/1995), Molière (1665/2017), and Tirso de Molina (1630/1990).

Despite the differences among the various versions of Don Juan characters, all of them share one common trait: the protagonist's insatiable sexual desire for women and pleasure derived from deceiving them. Undoubtedly, one of the most shocking aspects about Don Juan is not only his elevated number of conquests, but also his misogyny and rejoicing in deceiving women and leaving them with no virtue or honor. Numerous psychologists, physicians, and literary critics have proposed a variety of explanations for the protagonist's behavior such as misogyny, hypersexuality, lack of empathy, a sexual inferiority complex, homosexuality, and the Oedipus complex (Marañón, 1924; Parr, 1994; Rank, 1924; Rodríguez Lafora, 1927; Rousset, 1985).

The protagonist of Tirso de Molina's *El burlador de Sevilla* (1630/1990), the Golden Age Spanish literary work in which the character of Don Juan first appeared, clearly illustrates his lack of empathy toward women and hypersexuality in act II when he says:

> Not for nothing am I
> Labelled the greatest trickster of Seville.
> My very favourite pastime, my delight's
> To trick a woman, steal away her honour,
> Deprive her of her treasured reputation. (p. 93)

The protagonist of this seventeenth-century play has often been described as a "trickster" due to the profuse lies and tricks he employs to seduce the various women in the play. The schemes used by the first Don Juan varied according to the class of his victims; whereas he used his charm, fame, and status as a nobleman to impress and court common women or he deliberately usurped the personalities of other men in order to have sexual intercourse with noblewomen. Don Juan's victims felt deceived, anguished, and disgraced once they found out they had been tricked.

In William Shakespeare's play *The Tragedy of Othello, the Moor of Venice* (1564–1947), romantic deception is not only associated with anguish but also with death. In the play, a military general named Othello believes that his wife, Desdemona, is deceiving him by having an affair with Cassio (a former soldier who Othello had recently promoted to Lieutenant). In reality, Othello is being deceived by Iago, the soldier who led Othello to believe that his wife was being unfaithful. Iago started the rumors due to the jealousy he felt after Othello promoted Cassio instead of him, although Iago considered himself to have been the better soldier. Additionally, Iago set a plan to provoke an argument between Cassio and other generals. When Othello found out that Cassio had been involved in the fight, he punished Cassio for the incident by demoting him from his rank.

Following this conflict and Cassio's demotion, Iago planted Desdemona's handkerchief among Cassio's belongings. This act aroused Othello's suspicions of an affair between Cassio and Desdemona, as Iago convinced Othello that Cassio received the handkerchief from his wife. As a result, Othello named Iago his lieutenant and confronted Desdemona, who he then strangled in their bed. When Emilia, Iago's wife and Desdemona's maidservant, discovered what happened, she exposed Iago, and Othello finally realized that Desdemona had not been unfaithful to him. In retaliation, Othello stabbed Iago but he did not grant him the mercy of ending his life, as he wanted Iago to suffer the rest of his life in pain. Iago never explained why he sabotaged Othello's marriage. In the end, Othello committed suicide. Interestingly, some literary critics have hypothesized that Iago's motives may have been based on homoerotic fantasies of wanting to take Othello away from Desdemona (Dugger, 2013).

Nobel Laureate Pablo Neruda authored several poems that describe the feelings often reported by people who experienced romantic deception. He wrote most of his poems in the first person which prompted critics to believe his poems were autobiographical. Neruda's own statements have reinforced this idea. Indeed, he is known to have said, "If you ask me what my poetry is, I must say, I don't know; but if you ask my poetry it will tell you who I am!" (Shull, 2009).

In the poem *A Song of Despair*, the poetic voice speaks about how his past love overwhelms him in the present moment ("The memory of you emerges from the night around me") (Neruda, 1924). Additionally, the referenced poem also highlights how the end of a relationship left the protagonist feeling deserted, like a shipwreck sank under the depths of the sea. Neruda (1924) alludes to still yearn for his past love, despite recognizing it as lost.

Neruda's most famous poem, *Tonight I Can Write (The Saddest Lines)*, highlights several issues that are suggestive of romantic deception: (1) the idea that the protagonist's loved one will be with a different person ("Another's. She will be another's / As she was before my kisses"), (2) the difficulties people experience while trying to forget the person they love ("Love is so short, forgetting is so long"), and (3) the statement indicating that the protagonist's loved one was responsible for causing him pain and his decision to no longer pursue her ("Though this be the last pain that she makes me suffer / and these the last verses that I write for her").

Although the poetic voice in Neruda's work appears to reference his own feelings of despair brought on by others, Neruda is believed to have been the one who deceived women by having extramarital affairs. While being married to his first wife, he began a relationship with his second wife, and while married to his second wife, he began an affair with his third wife. It has been suggested that his first book, *The Captain's Verses*, was first published anonymously in 1952 to hide from his first wife his feelings for Urrutia (Neruda's lover at the time and second wife) (Cruz, 2015). Neruda referred to Urrutia as "the one with the fire/of an unchained meteor" and says in a love sonnet included in this book, "I love you as one loves certain obscure things, secretly, between the shadow and the soul" (Neruda, 1972). Neruda built a house for Urrutia and secretly lived with her for three years before divorcing his first wife (Cruz, 2015). As portrayed in the motion picture *Neruda*, he was also known for engaging in casual sexual encounters throughout his life (Larrain, 2016).

The second category of romantic deception includes classic portrayals of romantic self-deception or of individuals who deceive themselves about the identity of their loved one, the nature of their relationship, or about the feelings of their significant other has for them. One of the most well-known examples of romantic self-deception in literature is the Spanish novel, *Don Quijote de La Mancha*, published between 1605 and 1615 by Miguel de Cervantes Saavedra, and is based on a story of romantic self-deception. Don

Quixote was regarded during the German Romantic period as the romantic hero par excellence due to his confrontation against the world that surrounded him and his constant fight for his own ideals. In fact, according to German romantics, Don Quixote not only exemplified the struggle between the real and the ideal, but he also suffered for his ideas, which made him the perfect romantic hero.

As a romantic hero and a knight, the actions of Don Quixote were guided by his love for Dulcinea, the princess that he worshiped, loved, and adored. However, Dulcinea was not a real character in the novel but the product of Alonso Quijano's madness (the novel's protagonist). Indeed, neither Don Quixote nor Dulcinea were real characters in the novel; they were the invention of Alonso Quijano as a result of his wish to become a knight.

At the beginning of the novel, Alonso Quijano, a poor, middle-aged nobleman, loses his mind after reading too many chivalry novels. He dreams of becoming a knight and leaving his house in search of adventures. He soon realizes that in order to fulfill his dreams, he needs to have a princess to love, a squire, and a horse. As a result, he gives himself a new identity (that of Don Quixote) and invents the character of Dulcinea based on his own neighbor Aldonza Lorenzo, a woman with whom he had secretly been in love. In addition, he finds a squire (his neighbor Sancho Panza) to accompany him during his adventures.

Neither Aldonza Lorenzo nor Dulcinea make a physical appearance at any time in the two parts of the novel. Instead, the reader learns of them through the narrator and other characters' dialogue. Aldonza Lorenzo is portrayed as a real, unremarkable woman; in contrast, Dulcinea is described as a sweet princess and a lady of extreme beauty with golden hair, rose cheeks, coral lips, and pearl teeth. Dulcinea exemplifies not only beauty but also goodness; she is perfection itself and the protagonist's ideal love according to Don Quixote.

In the famous episode of the Montesinos Cave, Don Quixote realizes that Dulcinea is not the woman he imagined. She is not the perfect princess or goodness that he worshiped, but rather an ordinary beggar. Don Quixote's romantic deception ultimately leads to his metaphorical fall and sad ending. Importantly, Alonso Quijano was not a stable person. He suffered from madness, melancholy, and nostalgia. He ignored reality and invented someone to love (someone who did not exist) likely to replace his true love: Aldonza Lorenzo. Don Quixote illustrates the use of self-deception and ignoring reality as strategies for avoiding the pain of being deceived or hurt by a loved one.

Dance, Music, and Visual Arts

The despair that people can experience after being deceived by a loved one has been portrayed by multiple art mediums, including dance, music, and plastic arts. For example, Tango, the Argentinian ballroom dance commonly

associated with sensuality and romanticism, often focuses on the anguish that can arise as a result of romantic deception. The lyrics of most Tango songs describe melancholic themes and are mainly inspired by "existential anguish" (Salmon, 1977). These themes include despair over being rejected or deceived by a woman. In *Por Una Cabeza*, Carlos Gardel, the world's most popular Tango singer, describes how a woman lied to him about her feelings leading him to question the meaning of his own life. Similarly, Gardel famously asks, "What to live for" if his beloved woman were to forget him.

Although recognized as a feminist icon, Frida Kahlo's artwork was often inspired by the anguish she experienced over being betrayed by her husband (Diego Rivera), who romantically deceived her multiple times over the course of their marriage. Indeed, just one year after marrying Frida, Diego had the first of numerous affairs. The most painful one for Frida was undoubtedly the affair Diego had with Cristina, Frida's younger sister. Two paintings in particular are believed to exhibit the pain that resulted from this betrayal: *A few Small Nips* and *The Wounded Table*. In the painting *A few Small Nips*, Frida painted the naked body of a woman who had been brutally stabbed by her husband. Frida indicated at the time that the subject of the painting was the homicide of a woman who had been unfaithful. The woman's husband and killer reportedly defended the attack in front of a judge by saying, "But I only gave her a few small nips!" (p. 68). Frida later told a friend that she felt compassion for the woman because she felt "murdered by life" due to Diego's affair with her sister (Richmond, 1994).

Frida's painting, *The Wounded Table*, disappeared in the 1950s but photographs of the painting remain. The setting of this painting resembles Da Vinci's *Last Supper*. Frida painted herself in the middle as the martyr, and her husband was painted as the Judas, or victimizer, hugging her (McEarney, 2017). Frida had affairs of her own, most famously with Leon Trotsky and Isadora Duncan. Although Diego tolerated and at times encouraged Frida's sexual adventures with women, he became furious when he found out about Frida's affairs with men. Diego and Frida divorced in 1939 but remarried a year later. After Frida's death in 1954, Diego said, "I realized that the most wonderful part of my life had been my love for Frida." Diego died just three years later of heart failure (Richmond, 1994).

As the described examples indicate, classic literature and the visual arts have a long history of showcasing romantic deception. Most classical and canonical literary works, as well as other types of artistic works, such as music, dance, and painting, have addressed the topic of romantic deception in different manners, which illustrate the importance that this broad subject has played in art throughout the centuries. In this section, we have been able to identify two major types of romantic deception in classical and canonical art: (1) interpersonal deception and (2) self-deception. In the next section, we will address these two categories again, but in light of contemporary portrayals of art, such as film, television, and literature.

Contemporary Portrayals of Romantic Deception

The classical portrayals of romantic deception have inspired contemporary portrayals in television, film, and literature. In this section, we discuss examples that mirror the two previously discussed categories: (1) interpersonal deception (e.g., "*donjuanesque*" characters), and (2) self-deception. Additionally, we explore stories of romantic deception that depict characters who conjointly deceive others about the nature of their relationship and literary works that deceive the public about the reality of relationships.

Film and Television

The Don Juan archetype can be found in many contemporary movies that portray the use of romantic deception and the protagonist as a womanizer, including *The Wolf of Wall Street* (2013), *Catch Me If You Can* (2002), and *Hitch* (2005). However, there are some important differences between the "*donjuanesque*" protagonists in these screenplays and portrayals of Don Juan in classic literature.

In *The Wolf of Wall Street*, directed by Martin Scorsese and starring Leonardo DiCaprio, the protagonist (Jordan Belfort) is a Wall Street stockbroker addicted to drugs and sex. He lies and tricks both women and men to achieve money, success, and personal satisfaction. He is a modern "trickster" or "deceiver" like Don Juan as he not only commits fraud and tricks rich men to increase his own wealth, but he frequently solicits sex and is unfaithful to his first and second wives. Despite the similarities with Don Juan, Belfort is distinct from this literary character because he does not seem to experience joy from deceiving women. Belfort is unfaithful to both of his wives, but he loves them. He suffers a major romantic disappointment when his second spouse files for divorce. When he learns that she wants to divorce him, he insults her, hits her, and tries to kidnap their daughter.

In the critically acclaimed movie *Catch Me If You Can*, directed by Steven Spielberg and also starring Leonardo DiCaprio, the protagonist is a con man who tricks people into believing that he is someone he is not. The main character of this screenplay, which is based on the life of Frank Abagnale, begins to deceive people when he is just a teenager. He poses as a substitute teacher in his French class and starts lying to people, including family members. Before reaching the age of 21, he pretends to be a pilot, a doctor, and an assistant prosecutor until he is caught by the Federal Bureau of Investigation (FBI). Throughout the movie, he lies and tricks several women, especially bankers. One of his *modus operandi* is to give them a necklace and praise them. On one occasion, he swindles a prostitute by giving her a false check, an action that resembles the episode in which Don Juan brags about not paying prostitutes in *The Trickster of Seville*. Despite the similarities with the literary character, Abagnale does not feel pleasure in deceiving women. He falls

in love with Brenda (Amy Adams) and wants to marry her. Unfortunately, the FBI attends their engagement party, and he is forced to flee one last time. Before leaving, he confesses his real identity to Brenda and begs her to meet him in Miami to escape from the authorities together. Instead, she goes to Miami with the FBI and Abagnale is deceived for the first time in his life.

In the American romantic comedy *Hitch*, directed by Andy Tennant and starring Will Smith, the main protagonist (Alex "Hitch" Hitchens) is a doctor who is an expert in seducing women. Similar to Don Juan, he is a master of the spoken word. Hitch teaches his clients to make women fall in love with them. However, unlike Don Juan, Hitch's schemes are guided by a moral code: He only helps men who really love women, feel hopeless because they have suffered romantically in the past, or are too shy. Although it can be argued that Hitch is a trickster because he helps other men win over women, this moral code leads him to reject a client after finding out that he only wants to have sex with a woman. The audience later learns that Hitch's profession and manipulative relationships with women are explained by the fact that he suffered a romantic deception when his college girlfriend cheated on him. The shock impeded him from having a committed relationship for many years until he meets Sara Melas (Eva Mendes) and falls in love with her. However, when Sara finds out about Hitch's occupation, she ends the relationship. At the end of the movie, they reconcile, which conveys the hopeful message to the audience that it is possible to move on from a romantic deception.

There are many characters in modern cinema that are influenced by Don Juan; though closer examination of these characters suggest that Hollywood has softened the "*donjuanesque*" protagonists by exploring their motivation for lying and by making these characters both the perpetrators and victims of romantic deception. In contrast, in the classic portrayals of Don Juan, the motivation behind the protagonists' behaviors is not presented. Their protagonists are always the perpetrators and never the victims of romantic deception, thus they portray the literary character of Don Juan as less human than his cinematographic "*donjuanesque*" counterparts.

An example of romantic self-deception is offered in the American movie, *He's Just Not That into You* (Barrymore, 2009) in which the main character, Gigi, repeatedly misreads men's ordinary behaviors and concludes they are romantically interested in her. She then becomes very upset if these men do not respond to her phone calls. During the movie, Gigi meets Alex, a bar owner who befriends her and explains to her that she has been deceiving herself by continuously misinterpreting men's mundane behaviors as signs of romantic interest.

In some cases, couples deceive others about the presence, absence, or type of romantic involvement they have because of financial benefits or potential losses, sexual needs, or political benefits. In the American political drama television series *House of Cards*, the main characters, Francis J. "Frank"

Underwood and Claire Underwood, have a marital relationship that is largely based on pragmatism and lust for power. They sleep in different bedrooms and know of each others' extramarital affairs. However, they take steps together to maintain the image of a perfect, faithful romantic relationship that is based on traditional values of marriage. The Underwoods are experts at deceiving others.

Contemporary Literature

A notable case of self-deception in contemporary literature is portrayed in the book, *Travesuras de la Niña Mala*, by Nobel Laureate Mario Vargas Llosa (2006). As opposed to Don Quixote, the protagonist's obsession with the woman he loves is very much driven by sexual desire. The protagonist in Vargas Llosa's novel is Ricardo, a Peruvian man, who falls in love with Lily, a woman described in the book as cruel, materialistic, and self-centered. Ricardo meets Lily in his adolescence but she soon disappears from his life. Years later, when Ricardo's dream of living in Paris becomes a reality, he finds Lily again and his feelings for her intensify. Over the years, Lily chooses multiple rich men over Ricardo and explicitly tells Ricardo that she does not love him while humiliating him with her words and actions. Despite Lily's actions, Ricardo deceives himself by holding onto the belief that she secretly loves him and truly wants to be with him. In a particularly dramatic part of their story, Lily is ill and in need of help, and Ricardo devotes himself fully to her recovery. Yet Lily abandons him as soon as she recovers, leading Ricardo feeling overwhelmed with sadness and he attempts suicide. As Ricardo's suicide attempt failed, he returns home to find out that Lily changed her mind and wants to have a relationship with him. Ricardo's pain turned into homicidal rage against Lily, which led him to say that only his or Lily's death would allow him to get rid of her. Lily calms his fury by saying, "If you want, you can kill me later, but now make love to me" (p. 284). Although Ricardo still experiences fury, his anger dissipates as the love and lust for her lead him to want to be with her regardless of the consequences. Ricardo's feelings toward Lily are portrayed in the book as obsessive and his suffering as unbearable. In several occasions, Ricardo reflects on the pain he has experienced and decides to avoid Lily. However, their story continues and despite countless betrayals by Lily, Ricardo's self-deception never ends.

Similarly, *Fifty Shades of Grey* (James, 2011) and *Twilight* (Meyer, 2005) perpetuate unrealistic beliefs about love, sex, and romance that have a real impact on viewers. Empirical studies have found that individuals who consume media that portrays relationships in an unrealistic manner are more likely to have dysfunctional relationship beliefs, which in turn have a negative impact on their real-world relationships (see Galician, 2004 for a review). *Fifty Shades of Grey* and *Twilight* illustrate several themes of romantic deception that promote dysfunctional relationship beliefs, including: (1) the

illusion of perfection; (2) lying to protect loved ones; and (3) deception as a means of asserting control.

In particular, Edward Cullen of *Twilight* and Christian Grey of *Fifty Shades of Grey* are masters at creating deceptive impressions. They are seen by others as handsome, wealthy, and unattainable; however, this illusion is used to conceal their dark natures. Carl Jung (1959) referred to this facade as the *persona*, which is used to repress the *shadow*, or dark and forbidden side of the self. The shadow is characterized by carnality, chaos, and exoticism, clearly illustrated through Edward's true nature as a vampire and Christian's sexual deviance. Exemplifying his ability to deceive others with his persona, Edward states, "I'm the world's most dangerous predator. Everything about me invites you in" (Meyer, 2005, p. 264). These men deceive all those around them by embodying the antithesis of their shadows and portraying themselves as highly disciplined, rational, and reputable.

Another theme of romantic deception portrayed in these novels is lying to protect loved ones. In *New Moon* (Meyer, 2006), Edward lies to his love Bella about wanting to end their relationship because he believes it is the only way to shield her from the danger he brings into her life. Similarly, in *Fifty Shades Freed* (James, 2012), Ana convinces her love Christian that she is leaving him so he is not suspicious when she withdraws 5 million dollars from their bank account to pay the ransom for his kidnapped sister. Overall, these novels suggest that deception is justified when it is used to protect loved ones.

A final theme of romantic deception common to these series is the use of deception to gain control. The relationship dynamics between the main characters of *Twilight* and *Fifty Shades of Grey* are similar in that there is a marked imbalance of power within the couples. There are several instances in which Bella and Ana must resort to deceiving their partners in order to assert control. For example, when Christian worries about Ana's safety, he forbids her from leaving the house while he is out of town. As a result, Ana deceives him by sneaking out of the house to meet her best friend for drinks (James, 2012). Similarly, when Edward forbids Bella from visiting her best friend, she rebels against his authority by sneaking away to her friend's house (Meyer, 2007). These women ultimately believe they must use deception to gain even a slight sense of autonomy in their decision-making.

In sum, *Twilight* and *Fifty Shades of Grey* portray romantic deception in a way that perpetuates dysfunctional relationship beliefs and behaviors. In particular, they construct inaccurate portrayals of romance by idealizing the shadow personas as something mysterious and exciting, by justifying well-intentioned lies, and by presenting deception as an effective strategy for manipulating power in the relationship.

In this section, we have described examples of both romantic interpersonal deception and self-deception in contemporary literature, film, and television. These examples illustrate how the topic of romantic deception has transcended time and remains a prominent theme for artists. The subsequent

section focuses on romantic deception in modern communication outlets and how these platforms are used not only to deceive others, but also to deceive oneself.

Romantic Deception Through Modern Communication Outlets

Just like classical and contemporary portrayals of romantic deception, modern communication outlets also serve as mediums of romantic deception. It is common practice for individuals to create overly favorable impressions of themselves and their relationships on social networking sites, such as Facebook (e.g., Barash, Ducheneaut, Isaacs, & Bellotti, 2010; Walter, Battiston, & Schweitzer, 2009). Studies have shown that a couple's activity on social networking sites influences not only others' perceptions of their relationship, but also their own perception of the relationship. For example, couples who post more pictures together or write endearing posts about their partners on social media sites are perceived as having higher quality relationships than couples whose relationships are less visible on social media (Emery, Muise, Dix, & Le, 2014). Furthermore, a longitudinal study found that couples who listed themselves as being "in a relationship" on Facebook were more likely to stay together in the long term than those who did not make their relationships public on social media (Toma & Choi, 2015).

Further, online dating services also afford users the unique opportunity to easily control the information they present to potential romantic partners, and it is quite common for users to create deceptive self-presentations (Ellison, Heino, & Gibbs, 2006). In fact, a study found that 81% of users lied about at least one personal attribute in their dating profiles (Toma, Hancock, & Ellison, 2008). Deception in online dating profiles is often subtler than blatantly providing false information, as users seek to balance their appeal to prospective partners with being authentic (Ellison, Hancock, & Toma, 2012). Therefore, users most often engage in selective self-presentation, or increasing one's attractiveness by embellishing positive attributes and omitting information that could create an unfavorable impression.

The use of selective self-presentation strategies in online dating is pervasive; however, the degree to which users misrepresent themselves is often minimal. Typically, information presented in users' online dating profiles deviates only slightly from the truth (e.g., adding an inch to one's actual height or subtracting a few pounds from one's actual weight) (Toma et al., 2008). Most users expect dating profiles to be at least somewhat misleading (Gibbs, Ellison, & Heino, 2006) and cite the risk of being deceived by other users as the biggest drawback to online dating (Ellison et al., 2012). Yet research has found that the more someone expects others to be dishonest in their profiles, the more they are dishonest in their own (Drouin, Miller, Wehle, & Hernandez, 2016). In sum, it appears that most users accept

selective self-presentation as an unfortunate but inevitable part of the online dating landscape. In line with the theories of impression management and social comparison (see Rosenberg & Egbert, 2011), modern communication outlets allow users to easily control the information they present about themselves.

EMPIRICAL LITERATURE ON ROMANTIC DECEPTION

Empirical studies in the field of psychology, for example, provide insight into the complex dynamics of romantic deception. In this section, we showcase psychological research on the neurobiological mechanisms of romantic deception, including a theoretical framework to describe the function of lying to a romantic partner and the consequences of doing so. Next, we highlight the many paradoxes of this phenomenon, particularly the discrepancy between valuing honesty in relationships and the high prevalence of romantic deception.

Neurobiological Mechanisms of Romantic Love and Deception

Falling in love is often described as an "emotional rollercoaster," which some theorize is caused by increases in dopamine during the early stages of romantic love (Cacioppo, Bianchi-Demicheli, Hatfield, & Rapson, 2012; Marazziti & Cassano, 2003). Elevated levels of dopamine are associated with euphoria, a pounding heart, intense preoccupation, sleeplessness, decreased appetite, and mood swings—all hallmark features of falling in love. Notably, this experience closely mirrors symptoms of bipolar disorder (Cousins, Butts, & Young, 2009), obsessive-compulsive disorder (Denys, Zohar, & Westenberg, 2003), and the psychoactive effects of amphetamine (Schmidt, Ritter, Sonsalla, Hanson, & Gibb, 1985), all of which involve elevations in dopamine. Another key feature of falling in love is that it is a stressful experience, as people face the anxiety that comes with an unfamiliar social contact and the fear that their feelings will not be reciprocated (Acevedo & Aron, 2009). In fact, people who are in early-stage love have higher levels of the stress response hormone cortisol than people who are not in a romantic relationship (Marazziti & Canale, 2004). Likewise, a landmark study by Aron et al. (2005) reported that early-stage, intense romantic love is associated with dopamine-rich reward and goal-specific representation regions of the brain, including the ventral tegmental and dorsal striatum areas. Aron and colleagues concluded that romantic love is better characterized as a complex emotion, such that it is a goal-oriented state that is associated with specific emotions, such as anxiety and euphoria.

Empirical studies across the field of psychology indicate that during the initial stages of romantic love, increases in dopamine and cortisol serve to regulate the feeling of trust. For instance, elevated dopamine produces

heightened focus and social memory, which may help individuals attend to and recall information that is critical to their mating decisions (Fisher, Aron, Mashek, Li, & Brown, 2002). On the contrary, cortisol is also associated with distrust, paranoia, and suspicion (Riedl & Javor, 2012), suggesting that it may help individuals avoid situations leading to romantic deception. Taken together, changes in dopamine and cortisol may be important in facing threats that individuals may incur as a result of being deceived by a romantic interest.

In contrast to the initial experience of falling in love, being in a stable, romantic relationship is associated with feelings of calm and security (O'Leary, Acevedo, Aron, Huddy, & Mashek, 2012). People who are in love have elevated levels of oxytocin (Fisher, 2000), a hormone also associated with increased trust (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005). This hormonal reaction may be an important mechanism for allowing people to accept the risks associated with establishing long-term romantic relationships. In addition, people in love tend to have decreased levels of serotonin, a neurotransmitter that reduces pain and stress, and appears to promote a sense of security and safety in the relationship (Fisher, 2006; Fisher et al., 2002). Overall, when individuals are in stable, loving relationships, hyperstimulation decreases, as there is less need to detect potential threats from other romantic interests (Marazziti & Cassano, 2003). As a result, there is increased activity in neurotransmitters that promote calm and security in order for the brain to operate economically.

As the described examples indicate, empirical studies in the field of psychology have a long history of addressing romantic deception. Most notably, numerous studies have attempted to define the brain systems of lust, romantic attraction, and attachment. The subsequent section addresses the nature of romantic deception across psychological, social, and interpersonal contexts.

Empirical Research on Deception Within Romantic Relationships

Deception in romantic relationships is uniquely and arguably much more complex than deception in other interpersonal contexts. The highly intimate nature of romantic relationships gives rise to conflicting desires: the desire for honesty and openness in the relationship, and the desire to have some degree of privacy (for a review, see Montgomery & Baxter, 2013). Complete transparency is often viewed as an ideal for relationships, but in reality, it is rarely practiced (Roggensack & Sillars, 2014). This section discusses the unique features of deception in romantic relationships from an empirical standpoint and highlights important paradoxes.

A key paradox in romantic deception is that it is quite common for people to lie to romantic partners, despite the adverse consequences. Lying is considered one of the most severe transgressions in romantic relationships (West, 2006), yet an overwhelming majority of people say that they have lied

to their partners (Knox, Schacht, Holt, & Turner, 1993). In fact, research indicates that people lie in over one-third of their interactions with romantic partners (DePaulo & Kashy, 1998). Although these findings suggest that romantic deception occurs frequently, it can have grave consequences. Researchers have found that when deception occurs within a romantic relationship, both partners experience less trust, satisfaction, and commitment, and the couple faces more frequent conflict (Cole, 2001; Finkenauer, Kerkhof, Righetti, & Branje, 2009).

A second paradox in romantic deception is that honesty and trust are often seen as essential parts of a healthy relationship (Metts, 1989), yet most couples choose to stay together even after their sense of trust is compromised by deception (McCornack & Levine, 1990a). One explanation is that "partners may endorse honesty as an ideal but apply different standards to situations when pragmatic concerns make it difficult to be completely honest" (Roggensack & Sillars, 2014, p. 179). Again, transparency is often considered ideal, but only 27% of people believe that absolute honesty is necessary in their relationships (Boon & McLeod, 2001). Moreover, there are some circumstances in which partners consider deception acceptable and even beneficial (e.g., lying to avoid upsetting a partner and deceiving to maintain a satisfying relationship). However, most lies told to romantic partners are used to protect the deceiver's self-interests or used for personal gain (DePaulo & Kashy, 1998).

How is it that couples maintain their relationships despite the damage that results from deceiving a partner or being deceived? One explanation for this paradox is that people often do not know when their partners are lying to them (Levine, Parks, & McCornack, 1999). People are biased toward believing that their partners are truthful; therefore, they are less accurate in detecting deception from a romantic partner than from a stranger. This truth bias is even more pronounced in highly involved relationships, such that greater commitment and intimacy is associated with lower accuracy in recognizing relationship deception. As romantic relationships progress, people ultimately develop greater trust in their partners' credibility and are more blinded to deceptions that occur in their relationships (Levine & McCornack, 1992). In contrast, when partners have higher levels of suspicion in their relationship, they are significantly better at discerning when their partner is being dishonest (McCornack & Levine, 1990b).

These findings beg the question: When it comes to romantic relationships, is ignorance bliss? Some scholars argue that deception is necessary for maintaining a stable, healthy relationship (e.g., Cole, 2006; Spitzberg & Cupach, 2009) because deception is often used to avoid conflicts that would do more damage to the relationship than would the truth (Peterson, 1996). Although being suspicious of a partner's trustworthiness increases the chances of catching a lie, distrust negatively impacts relationship well-being and the ability to effectively manage conflict (Kim et al., 2015). In sum, deception can be

both helpful and harmful to romantic relationships, though the likelihood of each outcome is highly unpredictable due to the complex nature of deception between romantic partners.

Conclusion

Romantic deception is a universal phenomenon that is innate to human relationships. The theme of romantic deception has pervaded all forms of artistic expression including literature, painting, poetry, music, film, and television. Moreover, it is a transcendent phenomenon, such that it has inspired both classic and contemporary works, and continues to evolve with recent developments in technology.

What has changed over the course of history is our perceptions of romantic deception and the resulting consequences of deceiving a romantic partner. In classical works of art, romantic deception was often associated with tragic consequences, such as committing suicide to find relief from the agony of being deceived, or being sentenced to death for deceiving a romantic partner. In contrast, contemporary portrayals of romantic deception illustrate that although it is not a positive experience, the consequences of deceiving or being deceived are not nearly as severe. Extant research findings align with contemporary portrayals of romantic deception as they suggest that romantic deception is often considered inconsequential (e.g., exaggerating one's height in an online dating profile). In fact, this body of research argues that romantic deception can even be considered adaptive in some contexts, such as regulating our sense of trust in potentially threatening situations, and for avoiding conflict with romantic partners.

Our understanding of romantic deception will continue to evolve, not only as scholars add to the empirical literature, but also as advances in technology revolutionize how we interact with romantic partners. It is likely, however, that due to its universality, romantic deception will remain an inspiration for artists and endure as a tool for manipulating the dynamics of intimate relationships.

References

Acevedo, B. P., & Aron, A. (2009). Does a long-term relationship kill romantic love? *Review of General Psychology, 13,* 59–65.

Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., & Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology, 94,* 327–337.

Barash, V., Ducheneaut, N., Isaacs, E., & Bellotti, V. (2010). *Faceplant: Impression (mis) management in Facebook status updates.* Palo Alto Research Center.

Barrymore, D. (Producer), & Kwapis, K. (Director). (2009). *He's not that into you* [Motion picture]. USA: New Line Cinema.

Becerra Suárez, C. (1997). *Mito y literatura (Estudio comparado de Don Juan)*. Vigo: Servicio de Publicacións da Universidade de Vigo.

Boon, S. D., & McLeod, B. A. (2001). Deception in romantic relationships: Subjective estimates of success at deceiving and attitudes toward deception. *Journal of Social and Personal Relationships, 18*, 463–476. https://doi.org/10.1177/0265407501184002.

Cacioppo, S., Bianchi-Demicheli, F., Hatfield, E., & Rapson, R. L. (2012). Social neuroscience of love. *Clinical Neuropsychiatry: Journal of Treatment Evaluation, 9*, 3–13.

Cervantes Saavedra, M. (1605/1615/1998). *Don Quijote de la Mancha*. Instituto Cervantes (Ed.). F. Rico. Centro Virtual Cervantes. Retrieved from http://cvc.cervantes.es/literatura/clasicos/quijote/.

Cole, T. (2001). Lying to the one you love: The use of deception in romantic relationships. *Journal of Social and Personal Relationships, 18,* 107–129. https://doi.org/10.1177/0265407501181005.

Cole, T. (2006). Intimacy, deception, truth and lies: The paradox of being close. *Entelechy, 7*. Retrieved from http://www.entelechyjournal.com/timcole.htm.

Cousins, D. A., Butts, K., & Young, A. H. (2009). The role of dopamine in bipolar disorder. *Bipolar Disorders, 11,* 787–806.

Cruz, A. (2015, July 10). The story behind Pablo Neruda's greatest love. *Vivala*. Retrieved from http://www.vivala.com/arts-culture/pablo-neruda-greatest-love/244.

de Molina, T. (1990). *El burlador de Sevilla* (A. Rodríguez López-Vázquez, Ed.). Madrid: Cátedra (Originally work published 1630).

Denys, D., Zohar, J., & Westenberg, H. G. (2003). The role of dopamine in obsessive-compulsive disorder: Preclinical and clinical evidence. *The Journal of Clinical Psychiatry, 65*, 11–17.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74,* 63–79.

Drouin, M., Miller, D., Wehle, S. J., & Hernandez, E. (2016). Why do people lie online? 'Because everyone lies on the Internet'. *Computers in Human Behavior, 64,* 134–142. https://doi.org/10.1016/j.chb.2016.06.052.

Dugger, C. J. (2013). *Othello: Homosocial desire and its conversion to homosexual desire*. CUNY Academic Works. http://academicworks.cuny.edu/cc_etds_theses/395.

Ellison, N. B., Hancock, J. T., & Toma, C. L. (2012). Profile as promise: A framework for conceptualizing veracity in online dating self-presentations. *New Media and Society, 14,* 45–62.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication, 11,* 415–441.

Emery, L. F., Muise, A., Dix, E. L., & Le, B. (2014). Can you tell that I'm in a relationship? Attachment and relationship visibility on Facebook. *Personality and Social Psychology Bulletin, 40*(11), 1466–1479. https://doi.org/10.1177/0146167214549944.

Finkenauer, C., Kerkhof, P., Righetti, F., & Branje, S. (2009). Living together apart: Perceived concealment as a signal of exclusion in marital relationships. *Personality and Social Psychology Bulletin, 35,* 1410–1422. https://doi.org/10.1177/0146167209339629.

Fisher, H. (2000). Lust, attraction, attachment: Biology and evolution of the three primary emotion systems for mating, reproduction, and parenting. *Journal of Sex Education and Therapy, 25,* 96–104.

Fisher, H. (2006, February). *Why do we love, why do we cheat* [Video file]. Retrieved from http://www.ted.com/talks/helen_fisher_tells_us_why_we_love_cheat/transcript?language=en.

Fisher, H. E., Aron, A., Mashek, D., Li, H., & Brown, L. L. (2002). Defining the brain systems of lust, romantic attraction, and attachment. *Archives of Sexual Behavior, 31,* 413–419.

Galician, M. L. (2004). *Sex, love, and romance in the mass media: Analysis and criticism of unrealistic portrayals and their influence.* Mahwah, NJ: Lawrence Erlbaum.

Gibbs, J. L., Ellison, N. B., & Heino, R. D. (2006). Self-presentation in online personals: The role of anticipated future interaction, self-disclosure, and perceived success in internet dating. *Communication Research, 33,* 152–177.

James, E. L. (2011). *Fifty shades of grey.* New York: Vintage.

James, E. L. (2012). *Fifty shades freed.* New York: Vintage.

Jung, C. G. (1959). Archetypes. *The collective unconscious* (pp. 53, 352). Princeton, NJ: Princeton University Press.

Kim, J. S., Weisberg, Y. J., Simpson, J. A., Oriña, M. M., Farrell, A. K., & Johnson, W. F. (2015). Ruining it for both of us: The disruptive role of low-trust partners on conflict resolution in romantic relationships. *Social Cognition, 33,* 520–542. https://doi.org/10.1521/soco.2015.33.5.520.

Knox, D., Schacht, C., Holt, J., & Turner, J. (1993). Sexual lies among university students. *College Student Journal, 27,* 269–272.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature, 435,* 673–676.

Larrain, J. D. D. (Producer), & Larrain, J. D. D. (Director). (2016). *Neruda* [Motion picture]. Chile: AZ Films.

Levine, T. R., & McCornack, S. A. (1992). Linking love and lies: A formal test of the McCornack and Parks model of deception detection. *Journal of Social and Personal Relationships, 9,* 143–154.

Levine, T. R., Park, H., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the 'veracity effect'. *Communication Monographs, 66,* 125–144. https://doi.org/10.1080/03637759909376468.

Marañón, G. (1924). Notas para la biología de Don Juan. *Revista de Occidente, II*(VII), 15–53.

Marazziti, D., & Canale, D. (2004). Hormonal changes when falling in love. *Psychoneuroendocrinology, 29,* 931–936.

Marazziti, D., & Cassano, G. B. (2003). The neurobiology of attraction. *Journal of Endocrinological Investigation, 26,* 58–61.

McCornack, S. A., & Levine, T. R. (1990a). When lies are uncovered: Emotional and relational outcomes of discovered deception. *Communication Monographs, 57,* 119–138. https://doi.org/10.1080/03637759009376190.

McCornack, S. A., & Levine, T. R. (1990b). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communication Monographs, 57,* 219–230.

McEarney, A. (2017, March 25). *What happened to Frida Kahlo's missing adultery painting?* Retrieved from https://www.thedailybeast.com/what-happened-to-frida-kahlos-missing-adultery-painting.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6,* 159–179. https://doi.org/10.1177/026540758900600202.

Meyer, S. (2005). *Twilight*. New York, NY: Little, Brown.

Meyer, S. (2006). *New moon*. New York, NY: Little, Brown.

Meyer, S. (2007). *Eclipse*. New York: Little, Brown.

Molière. (2017). *Don Juan o el festín de piedra* (C. Manzano, Trans.). Barcelona: Penguin Random House (Original work published 1665).

Montgomery, B. M., & Baxter, L. A. (2013). *Dialectical approaches to studying personal relationships*. New York: Psychology Press. https://doi.org/10.4324/9781315045016.

Neruda, P. (1924). *Twenty love poems and a song of despair*. London, UK: Penguin.

Neruda, P. (1972). *The captain's verses*. New York, NY: New Directions Publishing.

O'Leary, K. D., Acevedo, B. P., Aron, A., Huddy, L., & Mashek, D. (2012). Is long-term love more than a rare phenomenon? If so, what are its correlates? *Social Psychological and Personality Science, 3,* 241–249.

Parr, J. A. (1994). Introducción. *El burlador de Sevilla y Convidado de piedra*. Binghamton, NY: Center for Medieval and Early Renaissance Studies State University of New York at Binghamton.

Peterson, C. (1996). Deception in intimate relationships. *International Journal of Psychology, 31*(6), 279–288. https://doi.org/10.1080/002075996401034.

Rank, O. (1924). *Die Don-Juan Gestalt*. Leipzig-Wien-Zürich: Internationaler Psychoanalytischer Verlag.

Richmond, R. (1994). *Frida Kahlo in Mexico*. Rohnert Park, CA: Pomegranate Artbooks.

Riedl, R., & Javor, A. (2012). The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging. *Journal of Neuroscience, Psychology, and Economics, 5,* 63–91.

Roggensack, K. E., & Sillars, A. (2014). Agreement and understanding about honesty and deception rules in romantic relationships. *Journal of Social and Personal Relationships, 31,* 178–199.

Rodríguez Lafora, G. (1927). *Don Juan, Los milagros y otros ensayos*. Madrid: Alianza.

Rosenberg, J., & Egbert, N. (2011). Online impression management: Personality traits and concerns for secondary goals as predictors of self-presentation tactics on Facebook. *Journal of Computer-Mediated Communication, 17*(1), 1–18. https://doi.org/10.1111/j.1083-6101.2011.01560.x.

Rousset, J. (1985). *El mito de Don Juan*. México, DF: Fondo de Cultura Económica.

Salmon, R. O. (1977). The Tango: Its origins and meaning. *Journal of Popular Culture, 10,* 859–866. https://doi.org/10.1111/j.0022-3840.1977.1004_859.x.

Schmidt, C. J., Ritter, J. K., Sonsalla, P. K., Hanson, G. R., & Gibb, J. W. (1985). Role of dopamine in the neurotoxic effects of methamphetamine. *Journal of Pharmacology and Experimental Therapeutics, 233,* 539–544.

Scorsese, M. (Producer) [and others], & Scorsese, M. (Director). (2013). *The wolf of wall street* [Motion picture]. USA: Red Granite Pictures, EMJAG Productions [and others].

Shakespeare, W. (1947). *The tragedy of Othello, the moor of Venice*. London: Oxford University Press (Original work published 1564).

Smith, W. (Producer) [and others], & Tennant A. (Director). (2005). *Hitch* [Motion picture]. USA: Overbrook Entertainment.

Shull, J. A. (2009). *Pablo Neruda: Passion, poetry and politics*. Berkeley Heights, NJ: Enslow Publishers.

Spielberg, S. (Producer) [and others], & Spielberg, S. (Director). (2002). *Catch me if you can* [Motion picture]. USA: Amblin Entertainment, Parkes/MacDonald Productions.

Spitzberg, B. H., & Cupach, W. R. (Eds.). (2009). *The dark side of interpersonal communication*. Routledge. https://doi.org/10.4324/9780203936849.

Toma, C. L., & Choi, M. (2015). The couple who Facebooks together, stays together: Facebook self-presentation and relationship longevity among college-aged dating couples. *Cyberpsychology, Behavior, and Social Networking, 18*(7), 367–372. https://doi.org/10.1089/cyber.2015.0060.

Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34*(8), 1023–1036. https://doi.org/10.1177/0146167208318067.

Vargas Llosa, M. (2006). *Las travesuras de la niña mala*. Mexico, DF: Santillana.

Walter, F. E., Battiston, S., & Schweitzer, F. (2009). Personalized and dynamic trust in social networks. In *Proceedings of the third ACM conference on Recommender systems—RecSys '09* (p. 197). New York, NY, USA: ACM Press. https://doi.org/10.1145/1639714.1639747.

West, A. E. (2006). *Relational standards: Rules and expectations in romantic relationships*. Doctoral thesis, Australian Catholic University. https://doi.org/10.4226/66/5a94b47d5e4bd.

Zorrilla, J. (1995). *Don Juan Tenorio* (A. Peña, Ed.). Madrid: Cátedra (Original work published 1844).

"Congratulations, Your Email Account Has Won You €1,000,000": Analyzing the Discourse Structures of Scam Emails

Innocent Chiluwa

A scam email is a type of *phishing*, which generally refers to attempts to defraud an email account owner by tricking them to disclosing their private or security information. According to Alsharnouby, Alaca, and Chiasson (2015), phishing is a criminal mechanism that employs social engineering and technical subterfuge in an attempt to steal consumers' personal identity data and financial account credentials. These types of emails involve asking prospective victims to disclose sensitive information such as their usernames, passwords, and/or credit or debit card numbers, often for malicious reasons. Phishing scams have led to identity theft, loss of sensitive intellectual property, and the loss of national security secrets (Hong, 2012).

Specifically, scam emails are unsolicited, often marked as spam, usually with warnings such as "be careful with this message. Many people marked similar messages as phishing scams, so this might contain unsafe content…" They are differentiated from general spam emails by the fact that they are crime-oriented and can result or have actually resulted in scams, where a victim is defrauded of his or her money (Chiluwa, 2015). This type of cybercrime has raised serious concern around the world, especially due to the high rate of loss reported by victims. Wright and Marett (2010) reported that phishing has also been a major problem for information system managers and users for several years. In 2008, for instance, it was estimated that phishing resulted in close

I. Chiluwa (✉)
Department of Languages and General Studies, Covenant University, Ota, Nigeria
e-mail: innocent.chiluwa@covenantuniversity.edu.ng

to \$50 billion in damages to US consumers and businesses. A report released by Ultrascan Advanced Global Investigations, an anti-money laundering and risk management association (2013), also shows that globally, losses arising from email scams totaled \$12.7 billion in 2013. According to the report, individuals from the US, UK, and India were the primary victims of the scams. A more recent report by the Australian Cybercrime Online Reporting Network (ACORN) indicates that 200,000 reports about email scams were received with a total loss of \$59 million to investment scams in Australia alone in 2016. Losses reported to *Scamwatch*, ACORN, and other scam disruption programs totaled \$299.8 million in 2016. With the growing popularity of electronic commerce, researchers have estimated losses to exceed US \$1 trillion yearly globally (Vishwanath, Herath, Chen, Wang, & Rao, 2011).

Most phishing email messages use social techniques, rather than technical tricks to fool their recipients (Hong, 2012). Very often, the phisher pretends to be a system administrator with an urgent message of warning about a new virus attack, urging the recipient to install an attached antivirus pack. If an unsuspecting recipient clicks on a suggested link, it directs them to a fraudulent website or tricks them into installing malware on their computers. Other forms of cyber-attacks operate the same way. The 2017 ransomware crisis that affected over 150 countries, starting from within UK hospitals, is a noteworthy and well-known example (Bodkin et al., 2017).

By applying social engineering tactics, phishers use the names of existing credible businesses (e.g., PayPal, eBay, Bank of America) and government institutions (e.g., Internal Revenue Service) as well as recent tragedies asking for donations (e.g., to earthquake survivors or *Katrina* victims). Some phishers claim to need funding for building orphanages, churches, or mosques. Invitations from phishers often invoke fear, threat, excitement, and/ or urgency to persuade people to respond (Vishwanath et al., 2011).

Types and Manifestations of Scam Emails

From Hong's (2012) categorization of scam emails, the "Nigerian 419 scam email," or "Nigerian 419 advance fee fraud," falls under what he called "fake phishing email," which according to him, has spread beyond email to include SMS, instant messaging and social networking sites. In Nigeria, scam emails are referred to "419 emails" or "yahoo-yahoo mail" (Chiluwa, 2009). "Yahoo-yahoo mail" is a Nigerian coinage for any suspicious email that seeks financial assistance or proposes a business opportunity. "419" is a section of the Nigerian Criminal Code that deals with financial fraud.

Other types of phishing include spear phishing or "whaling," where a receiver in this case, a management staff or a high military personnel receives an invitation to attend a dinner or party and is asked to click on a link to confirm his or her participation. Phishing also comes in the form of setting up fake websites or monetizing stolen information taken from stolen credit cards (Hong, 2012).

Scam emails often appear as investment openings or business proposals (Chiluwa, 2010). Such proposals may include an invitation to a recipient to provide his or her bank details for a possible money transfer; the recipient is told they will receive about 40% of the transferred money as compensation afterward. A similar email may invite a recipient to claim a huge sum of money deposited in a dormant bank account in West Africa (usually at Ouagadougou, in Burkina Faso) in the name of the recipient. Some scam emails also invite the recipient to act as a next of kin to an unknown person who has deceased and to claim his or her wealth. Such emails come in the form of bequeathing a large sum of money, where a supposed benefactor had died suddenly and left his or her material fortune for the email account owner (see Sample 1 for an example). Burrell (2008) argued that scamming is "a problematic empowerment" because scammers seek to unite their own interests with the self-interest of their global audience or targets. And "while successful internet scams may transfer wealth and, therefore some degree of power to scammers, they come at a broader cost to national and ethnic reputations and representations. Ultimately, they affect access of the larger society to legitimate venues for self-improvement" (p. 27). Generally, a recipient of a scam email is required to disclose some personal or security information to the writer in order to proceed with the transaction.

> **Sample 1**. Mr. Mavis d. Maxwell <www.@chime.ocn.ne.jp>
> Attention: beneficiary,
>
> Your inheritance payment of seven million five hundred thousand United State dollars ($7,500,000,00) is now approve to send to you through money gram transfer agency, due to your inability to pay airport clearance, so we have to call back your payment and deposit it to money gram agent. You're to contact the manager to start your daily payment immediately, and you will receive it in any money gram agent around you. The maximum amount you will receive daily is $980,000 us dollar, you cannot receive more than $980,000 daily.
>
> Forward your full information to the money gram agent to this information: (ifeanyihno.karl.raphael@gmail.com) phone: +229-987-204-78, his name: Mr. Raphael Ifeanyihno.

Further, other fraudulent emails announce charity donations to recipients and/or claim that the recipient has won a lottery. For example, email recipients have been told that they have won millions of Euros from a supposed random selection (see Sample 2).

> **Sample 2**. "You have won €1,000,000.00"
> Dear Lucky Beneficiary,
>
> You have been selected to receive the sum of "€1,000,000.00" as charity donations/aid from the Qatar Foundation, on the 20th of June 2016. Contact Mr. Rashid Al-Naimi through e-mail for more information: rashidalnai@gmail.com.

Yours Sincerely,
Mr. Rashid Al-Naimi.

IMPORTANT: If you receive this message in your spam or junk folder, it's due to your network provider, kindly move the message to inbox folder and reply back for the donation award claims.

These forms of online deception, or "digital lies" (Heyd, 2008), are a relatively new area of study. While studies of phishing attacks have increased in the last couple of years, studies that focus on scam mail or advanced fee fraud, which is the focus of this chapter are still quite few. In the next section, I examine the sources of scam emails as well as approaches and theoretical perspectives that have been applied to phishing and scam emails.

Sources of Scam Emails

The sources (or writers) of scam emails are generally unknown, though some authors (e.g., Heyd, 2008) have suggested that they are written by Nigerians. Zook (2007) also argued that advance fee fraud "has strong historic ties to Nigeria...and operates via a globally dispersed network that contains a clear agglomeration of activity in West Africa" (p. 65). According to a report by the *Ultrascan Advanced Global Investigations*, about 85,000 of the perpetrators are said to come from the "Nigerian Diaspora," residing in 69 countries. This could be what Blommaert (2005) meant when he suggested that the writers come from the "periphery of the world" writing to addressees in the "core countries of the world system" (p. 2).

While admitting the complicity of Africans in Internet scamming activities, Burrell (2008) lamented that the deployment of fictional narratives depicting political turmoil, corruption, violence, poverty, and personal tragedy set in African nations indeed stereotyped representation of Africa and Africans. Hence, Chiluwa (2009) has argued that the writers might not exclusively be Nigerians or Africans, because some of the addresses shown on the emails are presumably from Europe, Asia, and the Middle East.

Hong (2012) argued that email scams thrive on human greed because scammers generally defraud their victims through some form of promise to enrich them. So, both the scammer and the scammed are to blame, irrespective of where they come from. Because of the prevalence of scam emails, Blommaert (2005), further argued that phishing email is a function of globalization, as the "lowest empirically observable level of globalization...that is distributed by the globalized channel par excellence – email" (p. 2). And this type of globalized communication succeeds not only because they have unique textual features, but also textual functions in which they try to articulate identities, intentions, and contexts.

Approaches to the Study of Phishing and Scam Emails

Research on phishing, including the Nigerian (419) scam, has attracted scholarly attention in the fields of linguistics, information technology, and communication (e.g., Chiluwa, 2009, 2010; Heyd, 2008; Mintz, 2002; Orasan & Krishnamurthy, 2002). Most studies of phishing discuss its general features and characteristics, and how to recognize and prevent phishing attacks (e.g., Hong, 2012). Wright and Marett (2010) applied interpersonal deception theory to study phishing susceptibility and identified some experiential and dispositional factors that may increase people's likelihood to complying with phishers' requests for personal information. In terms of experience, the study argues that Internet users with high computer self-efficacy (CSE) are most likely to effectively handle online threats and secure their privacy; while those with low CSE are more prone to fall victims to online deception. The study further argues that a person's disposition to trust and willingness to believe and depend on others is an important factor toward phishing outcomes. Since deception depends on a high degree of trust, people who report a high disposition to trust are most likely to fall prey to requests for information from strangers.

Alsharnouby et al. (2015) examined strategies for combating phishing attacks. Using eye tracking usability methods, their study tested whether improved browser security indicators and increased awareness of phishing could lead to users' improved ability to protect themselves against phishing attacks. The study found that users successfully detected only 53% of phishing websites even when primed to identify them. Further, users generally spent very little time looking at security indicators compared to website content when making assessments of visual cues. Alsharnouby et al. conclude that "users' general technical proficiency does not correlate with improved detection score" (p. 69).

Vishwanath et al. (2011), also concerned with phishing susceptibility, presented an integrated information processing model, which showed that "most phishing emails are peripherally processed and individuals make decisions based on simple cues embedded in the emails" (p. 576). The study further argues that habitual patterns of media use and high level of email load have a strong influence on individuals' likelihood to be phished and suggests that CSE is significantly helpful for managing phishing susceptibility. This is similar to the finding by Alsharnouby et al.

Studies that examine the style and textual features of email scams (Chiluwa, 2009) or email fraud (Blommaert, 2005; Blommaert & Omoniyi, 2006) are also quite few. Among the very first sociolinguistic studies of email scams investigated English indexicality and fraud in email spam messages (Blommaert, 2005). The study analyzed the level of English competence of the writers and concluded that the writers demonstrated a "grassroots" level of English, which did not match their digital literacy. In other words,

scammers demonstrated better computer skills than the use of English. Blommaert's study further identified generic features of email fraud with the different indexical information they presented. The study called for more linguistic and generic studies of this genre of online communication. Following Blommaert (2005), Blommaert and Omoniyi (2006) argued that the authors of email fraud demonstrated technical competence to explore the opportunities that the global email systems offered to them. However, the writers lacked linguistic competence in order to produce messages that are appropriate to the projected identities and relationships in the proposed transactions. Barron (2006), however, carried out a macro-textual analysis of spam emails in order to investigate their promotional functions, and in particular, those associated with the promotion of medical supplies. The study showed that the emails were characterized by "obligatory moves" that consisted of "persuasive communicative purpose in the specific rhetorical context in which spam mail functions" (p. 100).

Heyd's (2008) book, *Email Hoaxes*, appears as the most rigorous linguistic study of all types of email scams that began to appear on the Internet from the late 1990s, including the identification of "virus hoaxes, giveaway hoaxes, charity hoaxes, urban legends and hoaxed hoaxes" (pp. 31–38). Heyd herself admitted that the distinction between terms such as *virus, spam*, and *hoax* was not clear-cut and might be confusing. With regard to the terminology used to discuss email hoaxes, she notes that "as a speech community" the field will negotiate "its names and expressions for an emerging technology" (p. 13). For example, she discusses scam emails as also being referred to as "Nigeria mail." Heyd's work was described as a genre study that set out to describe the various types of email scams as well as their structural and discourse features. Analyses of data were based on a linguistic and discourse qualitative methodology that described the forms of email hoaxes, their pragmatic contents, and communicative purposes.

Soon after, Chiluwa (2009) applied linguistic pragmatics to analyze the discursive-pragmatic contents and structures of scam emails. The study showed that writers applied sociocultural greeting formulas, reassurance, and confidence building as well as action-prompting strategies to sustain the interest of receivers. The study argued that economic hardship might have driven perpetrators of this crime into creative and complex ways of improving their living conditions.

Data Samples

Samples of scam emails in the current chapter were actual emails sent to recipients. Samples were collected from the author's inboxes and that of his students and colleagues between 2009 and 2017. Some members of the author's university community who received such emails were also asked to forward them to the author for study purposes. A point to note here is that the data are used unedited and appear in their original form. Most of

the samples exemplify grammatical errors and mistakes with figures, such as quoted sums of money in words that do not correspond with their figure equivalents (e.g., Sample 1).

Textual Features of Scam Emails

Generally, scam emails are letters written as narratives; thus, their textual structure consists of the usual introduction, content, and conclusion. The introduction includes the opening in the form of a greeting and introductory note about the writer. The greeting formula or salutation is usually "sir," "dear sir," "dear friend," or "hello". Some of the authors do not introduce themselves; they simply begin with a greeting and proceed to introduce the "business." The examples below include the opening and each author's introduction:

> **Sample 3**. Sir, with respect and humility am writing you this proposal letter which I hope may be of interest to you. I am Mrs. Juliet Annita Khubeka, am the wife of late Mr. John Khubeka who was murdered by the Zimbabwean veterans and irate black people along with other members of the family…

> **Sample 4**. Hello Friend…. I have a good business proposal which i want to let you know about…

> **Sample 5**. Dear Sir/ Madam,
> My name is DR. Ali Tarhouni, from Tripoli– Libya, I am former Minister for Oil and Finance on the National Transitional Council, but presently I'm here in Burkina Faso (West Africa) …

> **Sample 6**. Dearly co-worker in Christ,
> May the Lord bless you and your family and all that you do over there in your Country for our Lord Jesus Christ…

Some of the scam emails express a religious tone as in Sample 6. The introduction of such emails often begins with a prayer as in the above example.

The narrative content of the emails explains the subject matter with some form of persuasive argument. The concluding part generally ends with a complimentary tone and sign-off with the writer's name, and sometimes includes a short explanatory note of reassurance. Some others (see Sample 2) may contain an "important" notice, where the recipient is advised to move the message to his or her inbox if it is already in their spam folder. The conclusion and sign-off are generally appealing in nature—some of them appearing as complimentary, and similar to those of business letters such as "yours sincerely," or "best regards" as in Samples 7, 8, and 9.

> **Sample 7**. Thank you.
> Yours sincerely,
> Mrs. Juliet Annita Khubeka.

Sample 8. Best Regards,
Dr. Ali Tarhouni

Sample 9. Regards,
Mr. Mavis D. Maxwell
Secretary UN Charity Development Agency

Sometimes, email scammers adopt professional titles such as "Dr." or "Professor" in order to lend creditability to the proposal. They also claim false job designations such as "Secretary, UN Charity Development Agency" or "Governor, Central Bank…" (e.g., Sample 9). Conclusions may be informal, and sometimes make an emotional appeal to the receiver, especially where there is a story of a sudden death. Some conclusions adopt religious tones in order to appeal to religious sentiments. About 90% of the samples in the data end with "regards," "best regards," or "God bless you."

Sample 10. Allah be with you
Regards
Dr. Ayesha Gaddafi

Sample 11. Your sister in the Lord,
God bless you and your family,
Mrs. Kate William.

Some earlier findings have established that the tone and style of scam emails generally resemble that of everyday interpersonal emails, which may make them difficult to distinguish (Orasan & Krishnamurthy, 2002). However, others have argued that the use of English language in scam emails is often unsophisticated and mediocre. According to Blommaert (2005), the language of scam emails is at the "grassroots" level of English. Therefore, scholars have surmised that writers of scam emails come from non-English speaking countries (e.g., Africa and Asia) (Chiluwa, 2015).

DISCOURSE STRUCTURES OF DECEPTIVE EMAILS

Discourse structures are the various ways writers present their arguments, including their narrative technique and authorial stance, as well as pragmatic patterns, propositions, and frames that have been used to persuade the reader. Unfortunately, some email account owners have been convinced by scam emails, and therefore unwittingly disclosed their bank details to phishers following seemingly genuine or emotionally moving narratives in the emails. This is why warning messages that follow suspicious emails appear as: "*be careful with this message. Similar messages were used to steal people's personal information. Unless you trust the sender, don't click links or reply with personal information…*" Some scholars have argued that this kind of deception thrives on human penchant for free money and the get-rich-quick syndrome

(Hong, 2012). The narrative patterns and techniques in the emails are discussed below, followed-up with a discussion of authorial positioning, argumentation, stance, and engagement.

Narrativity

Most scam emails are business-like, brief, and straightforward, though often incoherent, and replete with grammatical errors, as highlighted above. They are generally written as formal announcements in supposed technical business English. The subject header is also businesslike with memo-like features (e.g., Sample 12). As a form of business discourse, the emails have three segments, namely the subject header, which includes, for example, information about a lottery sponsoring agency or lottery winner announcement; followed by the detailed message; and the closing. The sign-off includes the name of the writer and his contact information. Sample 12 shows an email sent to a receiver that is said to have won 500,000 Euros. The two forms of scams (e.g., donation and lottery winning) are basically informative; the donation message was accompanied with a video to validate the author's claim. In the end, the recipient of the email is expected to supply their personal contact information.

> **Sample 12. 2009 E-LOTTERY BONANZA**: CONFIRM YOUR PRIZE AWARD ASAP.
> Thursday, January 29, 2009 9:20 AM
> From: "mohn@eircom.net" <mohn@eircom.net>
> Add sender to Contacts
> To: markvanbossen2010@yahoo.com
> 2009 E-LOTTERY BONANZA: CONFIRM YOUR PRIZE AWARD ASAP.
> Sponsor Loterij~Awards International Amsterdam, The Netherlands.
> Lottery Claims/Service Department.
>
> Sir/Madam,
> This is to inform you of the Lottery Result of The Sponsor Loterij~ Awards International, which was held on the Monday 26th January 2009, with the aid of the E-Ballotting System.
>
> Your e-mail address attached to E-Ticket Number: 34–11–27–51 (4–82), with Reference Number: NW-417–8090–08 and Batch Number: AMSNL2ND-0110 drew a prize of 500,000.00 (Five Hundred Thousand Euros). This lucky draw came first in the 2nd Category of the Sweepstake by an e-ballot draw from over 50,000,000 e-mail addresses (personal and corporate e-mail addresses). To receive your won prize you are adviced to contact our appointed claim agent for you below who will facilitate the process of the claim of your won-prize. Note That because of the amount of winners in the different categories our management has out sourced the claim procedures and processes to the Government Accredited Agent below to assist all prize award winners complete all claims procedures so that the paying bank can then effect payment.

Contact your Appointed/Accredited Claim Agent:
Mark Van Bossen.
Tel/Fax: 0031–847–304–770
Email: kingzefinancialbv@aol.nl
markvanbossen2010@yahoo.com
Accept our heartfelt congratulations on your lottery prize winning.
Paula Van Mohn Lottery Co-coordinator

Sample 13. Donation for you
The message from me is not a mistake or fake Abdel Wassim is my name, Am from one of the wealthy family of the Middle East in Qatar we have investment all over the world. I am a charitable person, now I am giving 1.5 million USD as a donation. My donation may sound unbelievable as i do not know you, But with the help of cash promotion companies you are selected on the basics of mobile #, email, humanitarian act, profession, income and skills.

This is real and not fake, WATCH ME HERE: https://www.youtube.com/watch?v=tpp2ARJ48wg; you will see a video of me which talks more about myself and my donation to people in terms of Cash, House and luxury car. Send your Full Name, Age, Country, and Phone Number. on my private email: abdel-wassim@aim.com.

Another form of the narrative technique is "tellability" (Chiluwa, 2009), where the writer tells his or her story or gives the full description of the so-called business. Here, the narrator tells an explanatory story of either an "abandoned fund," which he or she intends to transfer to the recipient's accounts (e.g., Sample 14); or a sad story of an embattled family. The story in Sample 15 is that of a man who was killed for supporting the opposition party against the government of Robert Mugabe of Zimbabwe. The family had managed to escape to South Africa and now wants to invest the family's money (i.e., $35.4 million) in the email recipient's country. Since the death of the writer's husband (the writer is the widow of the dead man), there has been a serious threat on their lives; the email account owner is asked to help transfer the money to their account for the purpose of investment. He or she would get about 30% of the total money as reward in the end. Many of the emails in this category make use of real names of people and well-known places in order to appear authentic. The writer attributes the death of her "husband" to his support of the Zimbabwean Movement for Democratic Change (MDC), led by the late Morgan Tsvangirai, a major opposition figure to the former President Mugabe. Tsvangirai was the Prime Minister of Zimbabwe from 2009 to 2013.

Sample 14. I have a business proposal in the tune of $15.3 m USD for you to handle with me. I have opportunity to transfer this abandon fund to your bank account in your country which belongs to our client. I am inviting you in this transaction where this money can be shared between us at ratio of 60/40% and help the needy around us don't be afraid of anything I am with you I will instruct you what you will do to maintain this fund. Please kindly contact me

with your information's if you are interested in this transaction for more details (salem.almuhannadi@outlook.com).

1. Your Full Name; 2. Your Address; 3. Your Country of Origin; 4. What do you do for living; 5. Your Age; 6. Gender; 7. Your ID card copy and telephone number for easy communication. I will be waiting for your response.
Regards,
Mr., Salem Al Muhannadi
Tel: +228 91 39 60 20

Sample 15. I am probably of the view that you are aware of the present situation in my country Zimbabwe sequel to the Land/Farm reform Act crisis in Zimbabwe. Before the death of my husband on the 2nd of February 2006, he was a Prosperous Tobacco farmer and the Former President of the Commercial Farmers Union and one of the Major Sponsors of the Opposition Party: Movement for Democratic Change (MDC) in Zimbabwe led by Mr. Morgan Tsvangirai. I am writing to you together with my family presently in South Africa where we have taken refuge.

I got your contact from online directory service and decided to write this letter to seek for your assistance. I write to solicit for your special assistance to invest our Family fund in your country through your assistance in view of the existing peace in your country. Before the death of my husband and the consistent threats to his life because of his outstanding and vocal position against the government of President Robert Mugabe, He moved out the sum of US$35.4 Million, (Thirty-Five Million, Four Hundred Thousand United States Dollars), to South Africa since it is the nearest neighboring country and deposited it in a security and Trust Company.

Since the death of my husband, there have been threat to our lives, but I and my only son Frank Numa Khubeka succeeded to work for our escape to South Africa as political asylum seekers since that was the only option left for my family. In view of this development, our position in South Africa do not permit us to normalize this fund for any meaningful business transaction in South Africa since we live in South Africa on a refugee status, and the financial policies of this country do not permit us, we are soliciting for your help as a foreign partner.

That is why I want this fund to be transferred into your Nominated account so that you will assist us to invest it in your country in any meaningful business venture. Hence, if you agree to assist us, we have two options for you, we can go into partnership with you and your family the moment this fund is moved or transferred out of South Africa or in the alternative, we offer you 25% of the money for your assistance, 5% will be for the possible expenses incurred in the process of this transaction, while 70% will be for our family investment in your country, through your assistance.

I want you to understand that this fund is purely a family fund / money. Money made after years of our family investment in farming but we will appreciate it if you will maintain a high level of confidentiality because I and my family do not want to be traced by the agents of the ruling party and the obvious implication of exposure. If you are interested in assisting us, kindly contact my son Frank Numa Khubeka by phone: +27-83-982-7265 and email:

franknumakhubeka1966@yahoo.com so that he will furnish you with more details on the process to facilitate the movement of this fund. Please, endeavor to call my son as soon as you send email, we need urgent and confidential response. For your notice, I have some medical issues and always in and out of hospital

The storyline is usually followed by an invitation to the receiver to take some action; this is the point where the receiver is told what to do, such as contacting the writer by email or filling out forms. In Sample 15, the receiver is asked to contact the son of the writer with a given phone number and email address.

Stance and Engagement

Writers of email scams generally adopt some level of stance to reinforce their arguments. Stance is the writer's disposition or positioning in relation to the reader of the message. Writers generally present their arguments and viewpoints about the topic being presented through a systematic use of linguistic and rhetorical devices that express their knowledge, confidence, evaluation, judgment, and commitment (Biber, 2006; Biber & Finnegan, 1989). This is referred to as *epistemic stance*, and various devices are used in this process. *Boosters*, words that express certainty, such as "certainly," "surely," or "obviously," and *hedges* like "probably" or "possibly" that reflect confidence or doubt, are examples of epistemic stance. Also, writers not only convey their viewpoints, but they also express their attitudes and feelings, referred to as *affective stance*. Ochs (1990) defined affective stance as a "socially recognised feeling, attitude, mood or degree of emotional intensity" (p. 2). Significantly, deceptive scam emails are perfect examples of the discursive medium for the expressions of both epistemic and affective stances. In many scam narratives, clever rhetoric is supported by forms of commitment and emotion.

Similar to taking a position in an argument, the writer also engages his or her reader in the argument. *Engagement* is the point where "writers acknowledge and connect to others, recognizing the presence of their readers, pulling them along with their argument, focusing their attention, acknowledging their uncertainties, including them as discourse participants and guiding them to interpretations" (Hyland, 2005, p. 176). In most cases, the writers of scam emails express a "textual voice" and deliberately "stamp their personal authority" in their messages (Hyland, 2005, p. 175). By doing so, the scammers sound bold and authoritative. This is common with the donation and lottery winning announcements—the scammers position themselves as benefactors, or humanitarians, or even as saviors. The recipients are made to appear as if they are simply lucky, which is why an email recipient is sometimes accused of acting slowly. In Sample 1, the "beneficiary" was accused of his "inability to pay airport clearance," and hence, the so-called inheritance payment of $7.5 million was called back and deposited with a "money gram agent." Thus, in order to continue to sound credible and convincing, the

scammer employs grammatical structures such as imperatives to communicate authority, certainty, and commitment. Examples from the data include: "confirm your prize award ASAP" (Sample 12); "I am with you, I will instruct you on what you will do to maintain this fund" (Sample 14); and "This is real and not fake, WATCH ME HERE: …you will see a video of me which talks more about myself and my donation to people in terms of cash, house and luxury car" (Sample 13).

The body of messages that tell stories appear more engaging in terms of their seeming modesty or "humility," in disguising their real intentions. Unlike boosters, the writers try to involve the reader in the proposal and pretend to show some respect to the opinion of the receiver. So, they apply hedges in the form of words or expressions like "perhaps," "probably," or "I am not sure if …" and "I hope this message meets you well…" Some scam emails include apologies for intruding into the privacy of the receiver. In Sample 16, for example, the writer begins with "my greeting to you and sorry if this message came to you as a surprise…" Then, the writer, using a humble tone, asks the recipient to help move funds into their account.

Email scams also contain emotive words or expressions that evoke pity by telling stories of death, terminal disease (e.g., lung cancer), or air disasters. Some of the emails actually sound like the writing of someone at the point of death. As part of the engagement strategies, they appeal to shared knowledge and familiar topics, such as social and economic conditions including poverty, suffering, or child abuse. Some evoke religious sentiments such as in Sample 17, where the writer claims to "have prayed to the Lord to bring me a helper." The writer (the so-called widow) needed "a very honest and God-fearing person who can withdraw" the sum of $15.5 million she inherited from her late husband and use it for "charity work" or "work of God." Forty percent of the money would go to the recipient since the writer had no children. The writer of Sample 18, also a widow, wants her funds (i.e., $2.560 million inherited from her late husband) to be used to support orphans and widows, and to build mosques and charitable homes. Similarly, the writer of Sample 16, also a widow, who is suffering from "lung cancer and Parkinson disease," is releasing the sum of $12.5 million inherited from her dead husband for a "charity project." The textual and thematic similarity of these three emails suggests the possibility that one person could have written them. The writers who present themselves as "widows" draw from the cultural assumption that widows are helpless, pitiable, and always deserving care and attention. There is also the religious belief that God constantly rewards those who support widows.

Sample 16. God bless you and your family,

My Greeting to you and sorry if this message came to you as a surprise, My name is Mrs. Kate William widow, I found your contact through my husband dater late Mr. John William,Who died in a motor accident, My husband and I have no child before he died.

I am presently admitted at the hospital suffering from a lung cancer and Parkinson diseases, I have some funds at security company here where I currently live now, I inherited from my late husband account the amount of $12,500,000 million Dollars, I wish to know if I can trust you to use the funds for charity project,

If you have a good character or a charitable organization or a person working for children from poor families and have the fear of God, Kindle get back to me for more details; contact email; (katewilliam711@gmail.com)

Your sister in Lord,

God bless you and your family,
Mrs. Kate William

Sample 17. Dear friend,

Calvary Greetings in the name of the LORD Almighty and Our LORD JESUS CHRIST the giver of every good thing. Good day and compliments of the seasons, i know this letter will definitely come to you as a huge surprise, but I implore you to take the time to go through it carefully as the decision you make will go off a long way to determine my future and continued existence. I am Mrs. Nadesh aging widow of 64 years old suffering from long time illness. I have some funds I inherited from my late husband, the sum of ($15,500,000.00 Million Dollars) and I needed a very honest and God fearing who can withdraw this money then use the funds for Charity works. I WISH TO GIVE THIS FUNDS TO YOU FOR CHARITY WORKS. I found your email address from the internet after honest prayers to the LORD to bring me a helper and i decided to contact you if you may be willing and interested to handle these trust funds in good faith before anything happens to me.

I accept this decision because I do not have any child who will inherit this money after I die. I want your urgent reply to me so that I will give you the deposit receipt which the SECURITY COMPANY issued to me as next of kin for immediate transfer of the money to your account in your country, to start the good work of God, I want you to use the 40/percent of the total amount to help yourself in doing the project. I am desperately in keen need of assistance and I have summoned up courage to contact you for this task, you must not fail me and the millions of the poor people in our todays WORLD. This is no stolen money and there are no dangers involved, 100% RISK FREE with full legal proof. Please if you would be able to use the funds for the Charity works kindly let me know immediately. I will appreciate your utmost confidentiality and trust in this matter to accomplish my heart desire, as I don't want anything that will jeopardize my last wish. Please kindly respond quickly for further details
Warmest Regards,
Mrs. Nadesh

Sample 18. Greeting To You I am Mrs. Hasna Ahmed am married to Late Mr. Ahmed, a widow and I want to make a donation of $2.560 two million five hundred and sixty. Usd only to help Orphans and Widows and mosque and Charitable home in your Country and I assumed that you will be able to receive this Fund and use it to my wished to the needy in your country and i am seriously ill please always remember me in daily prayers because i don't know when

it will end. Reply back to me immediately for more details about this fund. Thanks, Mrs. Hasna Ahmed

Because of the highly suspicious nature of scam emails, and the fact that the claims in the emails are often contradictory—sometimes causing fear, anxiety, and disbelief—the writer endeavors to apply some expression such as "don't be afraid" or "this is not a scam," to reduce the uncertainty of the receiver. These types of messages are often found toward the end of scam emails.

Many of the scam emails examined in the current analysis end with an offer of confidentiality and the instruction to "reply," "contact" the writer, or "act fast." In some of the fake business proposal emails, the receiver is told to keep the proposal confidential. And in some, the subject header reads: "Top Secret." Scam emails also sometimes conclude with statements such as, "I will appreciate your utmost confidentiality and trust in this matter to accomplish my heart desire, as I don't want anything that will jeopardize my last wish" (Sample 17) or "please keep this proposal as a top secret and delete if you are not interested." Ironically, many of the emails ask for sincerity, trust, and confidence from the addressee. The request for confidentiality in these cases may be intended so that receivers do not report them to the police, but perhaps more purposefully to create a false sense of confidence in the receiver who is indirectly being persuaded to invest in the "business" without involving anyone else (Chiluwa, 2009). These procedures also imply that the transaction can proceed without any formal legal procedures or negotiations. However, the creation of a false impression that the transaction depends on trust and confidentiality is mere deception.

CONCLUSION

Significantly, contrary to Heyd's (2008) prediction that email scams might be non-existent in the future, Chiluwa (2009) has argued that global economic challenges might continue to promote the increase of cybercrime and fraudulent practices on the Internet. More recently, some studies have observed that advance fee fraud and deceptive emails are not declining at all (e.g., Dobovsek, Lamberger, & Slak, 2013). As more people become aware of phishing and scam emails, others still fall victim. I call upon the global academic community for continued study of deceptive messaging in mediated contexts, and also to warn the general public of dangers of online scams. This area of study presents key opportunities to advance our conceptual and theoretical understanding of deception, and to work pragmatically toward helping others avoid becoming victims to fraudulent, deceptive messages.

REFERENCES

Alsharnouby, M., Alaca, F., & Chiasson, S. (2015). Why phishing still works: User strategies for combating phishing attacks. *International Journal of Human-Computer Studies, 82,* 69–82.

Barron, A. (2006). Understanding spam: A macrotextual analysis. *Journal of Pragmatics, 38*(6), 880–904.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.

Biber, D., & Finnegan, R. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text, 9*, 93–124.

Blommaert, J. (2005). Making millions: English, indexicality and fraud. *Working Papers in Urban Language & Literacies, 29*, 1–24.

Blommaert, J., & Omoniyi, I. (2006). Email fraud: Language, technology and the indexicals of globalisation. *Social Semiotics, 16*(4), 573–605.

Bodkin, H., Henderson, B., Donnelly, L., Mendick, R., Farmer, B., & Graham, C. (2017, May 13). Government under pressure after NHS crippled in global cyber attack as weekend of chaos looms. *The Telegraph*. Retrieved from https://www.telegraph.co.uk/news/2017/05/12/nhs-hit-major-cyber-attack-hackers-demanding-ransom/.

Burrell, J. (2008). Problematic empowerment: West African internet scams as strategic misrepresentation. *Information Technologies and International Development, 4*(4), 15–30.

Chiluwa, I. (2009). The discourse of digital deceptions and '419' emails. *Discourse Studies, 11*(6), 635–660.

Chiluwa, I. (2010). The pragmatics of hoax email business proposals. *Linguistik Online, 43*, 3–17.

Chiluwa, I. (2015). Email fraud. In K. Tracy, C. Ilie, & T. Sandel (Eds.), *The international encyclopedia of language and social interaction*. Boston: Wiley.

Donovsek, B., Lamberger, I., & Slak, B. (2013). Advance fee frauds messages—Nondeclining trend. *Journal of Money Laundering Control, 16*(3), 209–230.

Heyd, T. (2008). *Email hoaxes*. Amsterdam: John Benjamins.

Hong, J. (2012). The current state of phishing attacks. *Communications of the ACM, 55*(1), 74–81.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173–192.

Mintz, A. (Ed.). (2002). *Web of deception: Misinformation on the internet*. Medford, MA: CyberAge Books.

Ochs, E. (1990). *Cultural universals in the acquisition of language: Keynote address*. Papers and reports on child language development. Stanford University, 29, 1–19.

Orasan, C., & Krishnamurthy, R. (2002, May 29–31). A corpus-based investigation of junk mails. *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Spain. Available at http://clg.wlv.ac.uk/papers/orasan-02b.pdf.

Ultrascan Advanced Global Investigations. (2013). Retrieved from http://www.ultrascan-agi.com/.

Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems, 51*, 576–586.

Wright, R. T., & Marett, K. (2010). The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. *Journal of Management Information Systems, 27*(1), 273–303.

Zook, M. (2007). Your urgent assistant is requested: The intersection of 419 spam and new networks of imagination. *Ethics, Places & Environment: A Journal of Philosophy & Geography, 10*(1), 65–88.

Contexts of Deceptive Communication: Court, Politics, and Propaganda

Credibility Assessment and Deception Detection in Courtrooms: Hazards and Challenges for Scholars and Legal Practitioners

Vincent Denault and Norah E. Dunbar

We all have lied, and we all have been lied to (DePaulo & Kashy, 1998; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Serota, Levine, & Boster, 2010). However, while consequences of undetected prosocial lies can be deemed innocuous, positive, or even ethical in some cases (Levine & Schweitzer, 2014, 2015; Wilthermuth, Newman, & Raj, 2015), the fallout from other types can be detrimental. For example, undetected lies from suspects can steer police investigations in the wrong direction and result in severe consequences. The perpetrator of a crime may potentially be released while an innocent individual is arrested. However, the failure to detect deception at the very last step toward justice, during trials, can arguably result in far worse consequences. In a criminal trial, for example, if a witness falsely testifies that a defendant committed a crime, the defendant could end up with a criminal record unfairly. If the crime is more

V. Denault (✉)
Department of Communication,
Université de Montréal, Montréal, QC, Canada
e-mail: vincent.denault@umontreal.ca

N. E. Dunbar
Department of Communication,
University of California, Santa Barbara, CA, USA
e-mail: ndunbar@comm.ucsb.edu

severe, the failure to detect deception may lead to wrongful imprisonment or the death penalty in countries where capital punishment is still carried out.

Obviously, deception in courtrooms is not limited to criminal trials. In civil and family trials, disastrous financial and human consequences can result from undetected deception. For example, if an entrepreneur were to falsely testify that a business partner made a mistake, the business partner's reputation could be ruined, and future contracts could be lost. If a father were to falsely testify that a mother is abusive, the mother could lose custody of her children. However, despite such significant consequences, the detection of deception in courtrooms, whether in criminal, civil, or family trials, remains underexplored when compared to the plethora of peer-reviewed publications on the detection of deception in investigative interviews and other contexts (Denault & Jupe, 2017; Fawcett, 2014; Granhag & Strömwall, 2004; Vrij, 2008). Such a situation places scholars and practitioners in a somewhat precarious position. If characteristics of trials (e.g., the rules of evidence and the judge's and jury's roles) are inadequately understood, scholars might provide inappropriate advice to legal practitioners (Denault & Jupe, 2018; Denault, Jupe, Rochat, & Dodier, 2017). Then, because of a lack of relevant and accessible knowledge, legal practitioners may unknowingly turn to false beliefs and inappropriate stereotypes about the detection of deception and even pseudoscientific notions (Denault, 2015; Denault & Dunbar, 2017). Considering that undetected deception in courtrooms severely jeopardizes the integrity of the judicial process, a cornerstone of democracy, this chapter aims to address an issue that is long overdue—the detection of deception in courtrooms.

To achieve this objective, we first offer an overview of deception in courtrooms of adversarial justice systems, that is, justice systems such as those in Canada and the US where parties to a dispute each have to present to a judge or jury their own evidence and arguments in support of their position (Sward, 1989). Second, using case law examples and empirical research on deception detection, we address the influence of false beliefs and inappropriate stereotypes on deception judgments made by judges or jurors. Third, we present limitations to the use of novel deception detection techniques during trials. We end this chapter with a call for scholars concerned with the search for truth and justice to give serious consideration to the study of deception detection in courtrooms of adversarial justice systems.

Deception in Courtrooms of Adversarial Justice Systems

During trials, when witnesses are testifying, some provide inaccurate information without the intent to mislead the court, notably as a result of memory loss and distortion (Lacy & Stark, 2013). However, others will do so

intentionally. In other words, for many reasons, some witnesses engage in deception, and some are more inclined to deceive than others (Fawcett, 2014; Loevy, 2006; Simon-Kerr, 2015; Slobogin, 1996; Vrij, 2007). For example, parties to a dispute or interested witnesses (i.e., witnesses who have something to win or lose from the outcome of the dispute) might be more inclined to deceive than selfless witnesses who have nothing to win or lose from their testimony (Fawcett, 2014; Farmer & Hancock, 2014). However, those witnesses might deceive for reasons such as intimidation or fear of retribution from the person against whom they testify (Browning, 2014; Buel, 2014; Connick & Davis, 1983). Parties to a dispute or interested witnesses might even be more inclined to deceive if the consequences of detection are lower than the consequences of losing their criminal, civil, or family trial (Fawcett, 2014). Deception strategies used during trials may also vary in consideration of such consequences.

Unlike vague and ambiguous statements from witnesses (i.e., the equivocation strategy) that lawyers have a duty to recognize (Buller, Burgoon, White, & Ebesu, 1994; Green, 2001; Kane, 2007), the omission strategy is likely one of the most prevalent and difficult to detect in courtrooms of adversarial justice systems. In civil and family trials, for example, parties to a dispute can have considerable discretion when deciding what evidence to present and how to present it. The law typically does not compel them to reveal evidence that could be detrimental to their case (Strier, 1994; Summers, 1999). Therefore, be it through equivocation or omission, parties to a dispute might engage in considerable deception without even breaking the law. However, other deception strategies expose witnesses to higher risks. Falsification, that is, the presentation of false information as true information, is one of them (Buller & Burgoon, 1994; Ekman, 1985; Schwelb, 1989). In fact, falsification can make witnesses most vulnerable to criminal charges of perjury, depending on the jurisdiction. For example, *Canada's Criminal Code* states that:

> …every one commits perjury who, with intent to mislead, makes before a person who is authorized by law to permit it to be made before him a false statement under oath or solemn affirmation, by affidavit, solemn declaration or deposition or orally, knowing that the statement is false. (1985, section 131)

Using the above examples of the entrepreneur's false testimony about the business partner and the father's false testimony about the mother, the entrepreneur and the father could be exposed to criminal charges of perjury, which could lead to jail time. However, it is worth noting that, in real-life court proceedings, criminal charges of perjury are arguably rare which, quite obviously, does nothing to deter witnesses from engaging in deception (Farmer & Hancock, 2014).

In addition to equivocation, omission, and falsification, three deception strategies often referred to within the deception literature (Buller & Burgoon, 1996; Buller et al., 1994; McCornack, 1992; Van Swol & Braun, 2014),

lawyers might use other sophisticated verbal and nonverbal deceptive practices to frame an image of the truth to strengthen the position of their client (Atkinson & Drew, 1979; Bell, Villalobos, & Davis, 2014; Bulow-Moller, 1991). For example, lawyers can exaggerate or minimize claims, take words out of context, make strategic disruptions as well as unsupported insinuations, and offer technically true but misleading statements. They can even attack the credibility of witnesses without any concern if their testimony is truthful or not (Denault & Jupe, 2018; Frankel, 1975; Galasinski, 2000; Rogers, Zeckhauser, Gino, Schweitzer, & Norton, 2017; Strier, 1994).

Although such sophisticated verbal and nonverbal deceptive practices are much less likely to result in criminal charges of perjury (Farmer & Hancock, 2014), their consequences should not be underestimated compared to that of equivocation, omission, and falsification. Deception strategies of all kinds can impede the search for truth of the judge or jury, and even if they are ethically questionable, lawyers use them within the boundaries of their legal and professional obligations (Fortune, Underwood, & Imwinkelried, 1996; Frankel, 1975). Let us not forget that the "truth is all too often sacrificed at the altar of legal victory, profoundly corrupting the reliability of the adversary system as a vehicle for truth" (Strier, 1994, p. 108).

DECEPTION JUDGMENTS MADE BY A JUDGE OR JURY

The interest in detecting deception has been around for thousands of years (Troville, 1939; Wise, 1845). However, the early days of modern empirical research on cues to detect deception date back to the 1960s when Ekman and Friesen (1969) suggested the *leakage hypothesis*. According to this theoretical statement, very quick movements from the face (i.e., micro-expressions) and body actions can help the detection of deception. Since then, however, scholarly interest in nonverbal behaviors has decreased, as many have been demonstrated to be faint and unreliable (e.g., DePaulo et al., 2003), and empirical research on verbal cues to detect deception gained momentum (Hauch, Blandón-Gitlin, Masip, & Sporer, 2015). However, given the absence of unambiguous nonverbal or verbal cues akin to Pinocchio's nose (DePaulo et al., 2003; Vrij, 2008) and low accuracy rates in correctly classifying liars and truth-tellers (Bond & DePaulo, 2008), scholars set a new course on deception detection research. Their focus shifted to interviewing approaches to elicit and enhance cues to deceit (Masip, 2017; Vrij & Granhag, 2012a). Procedures for investigative interviews, the very first step toward justice, were developed but the detection of deception during criminal, civil, and family trials where final and enforceable court decisions are made has been overlooked (Denault & Jupe, 2017, 2018; Fawcett, 2014; Granhag & Strömwall, 2004; Vrij, 2008). While there are unquestionably many ways of dealing with this lack of knowledge, perhaps one of them is to take a closer look at a central feature of trials, the credibility of witnesses,

more so considering that credibility is the "largest determinant of a deception judgment" (Bond & DePaulo, 2008, p. 487).

The Credibility Assessment of Witnesses

In adversarial justice systems, the facts of the case are typically presented by witnesses during their testimony. Subsequently, the law is applied to the narrative established by the trier of fact, that is, a judge in a bench trial or jurors in a jury trial, depending on the jurisdiction as well as the nature of the case (Bell, 2013; Griffin, 2013; Paciocco, 2010). However, while some facts can be agreed upon or undisputed, others can be contested. In such a situation, the credibility of witnesses, that is, the characteristic subjectively attributed to witnesses that makes them believable and convincing (Brodsky & Pivovarova, 2016; Doyon, 1999; Köhnken, 1989), will influence the decision of the judge or jury to accept one witness's words over another. This influence can be particularly important when no other evidence proves or disproves the testimony of witnesses (Bond & DePaulo, 2008; Bell, 2013; Porter, Campbell, Birt, & Woodworth, 2003).

While the rules of evidence and the judge's and jury's roles vary to a greater or lesser extent from one country to another, Canada's highest court offers a prime illustration of how the credibility of witnesses can be assessed in courtrooms of adversarial justice systems. According to the Supreme Court of Canada, while it is "an issue that pervades most trials, and at its broadest may amount to a decision on guilt or innocence" (*R. v. Handy*, 2002, p. 951), the credibility assessment of witnesses should not be subject to a fixed operation (*Vetrovec v. The Queen*, 1982). In addition to "the significant pauses in the responses, the changes in facial expression, the looks of anger, confusion and concern" (*P. (D.) v. S. (C.)*, 1993, p. 192) of witnesses, the judge or jury can consider numerous human characteristics:

> The general integrity and intelligence of the witness, his powers to observe, his capacity to remember and his accuracy in statement are important. It is also important to determine whether he is honestly endeavouring [*sic*] to tell the truth, whether he is sincere and frank or whether he is biassed [*sic*], reticent and evasive. All these questions and others may be answered from the observation of the witness' general conduct and demeanour in determining the question of credibility. (*White v. The King*, 1947, p. 272)

In other words, credibility assessment "must always be the product of the judge or jury's view of the diverse ingredients it has perceived at trial, combined with experience, logic and an intuitive sense of the matter" (*R. v. Marquard*, 1993, p. 248). As long as inappropriate stereotypes are not used, the trier of fact can "use common sense and wisdom gained from personal experience" (*R. v. S. (R.D.)*, 1997, p. 537), a feature echoing that of

other adversarial justice systems, including the US (Bennett, 2015; Hutchins, 2014).

However, while the idea that credibility assessment should not be subjected to a fixed operation conforms with empirical research showing the multidimensional nature of credibility (Buller & Burgoon, 1996; Pornpitakpan, 2004; Rieh & Danielson, 2007), the emphasis placed on common sense and wisdom raises questions, in both bench trials and jury trials (Norris & Edwardh, 1995). While one could expect legal practitioners to be expert lie catchers, they often hold false beliefs and inappropriate stereotypes and are not better at catching lies than laypersons (Bond & DePaulo, 2008; Denault, 2015; Porter & ten Brinke, 2009; Strömwall & Granhag, 2003; Vrij, 2008). Therefore, if judges do not receive adequate training to understand human behavior and jurors do not receive sufficient instruction to mitigate false beliefs and inappropriate stereotypes, their credibility assessment of witnesses could easily be distorted, more so if they are instructed to rely on their own judgment to determine whether or not someone is trustworthy. However, such distortions will not easily be corrected on appeal because credibility is an issue to be decided by the trier of fact only considering the advantage of hearing and seeing the witnesses, which is denied to appellate courts (*R. v. Béland*, 1987; *R. v. Brooks*, 2000; *R. v. François*, 1994; *R. v. W. (R.)*, 1992). Therefore, court decisions solely based on issues of credibility will typically become final and enforceable (*P. (D.) v. S. (C.)*, 1993; *R. v. Gagnon*, 2006).

However, the importance attached to demeanor has raised and still raises questions. Police officers, lawyers, and judges hold numerous false beliefs and inappropriate stereotypes which can affect their inference as to whether a person is lying or telling the truth (Denault, 2015, 2017; Strömwall & Granhag, 2003; Vrij, 2008). Furthermore, one could argue that the problem is not the use of demeanor during trials per se, but the use of demeanor with the very specific erroneous idea that demeanor offers objective cues to determine the quality of a witness (Burgoon, Blair, & Strom, 2008; Denault & Dunbar, 2017; Levine, 2010; Levine et al., 2011; Pryor & Buchanan, 1984; ten Brinke & Porter, 2013).

Unfortunately, despite the complexity of the credibility assessment of witnesses made by judges or jurors, courts from Canada and the US typically prohibit expert testimony on that issue because the expert would usurp the function of the trier of fact (Friedland, 1989; Norris & Edwardh, 1995; Porter & ten Brinke, 2009). This position is in keeping with the US Supreme Court according to which jurors "are presumed to be fitted for it by their natural intelligence and their practical knowledge of men and the ways of men" (*Aetna Life Ins. Co. v. Ward*, 1891, p. 88; *United States v. Scheffer*, 1998, p. 313). However, while the credibility assessment of witnesses in such a way has been met with severe criticism (Denault, 2015; Morrison & Comeau, 2002; Porter & ten Brinke, 2009; Porter, ten Brinke, & Gustaw, 2010; ten Brinke & Porter, 2013), how these principles

are, in practice, implemented by judges or jurors is not without shortcomings either. In fact, the assumption that credibility assessment "is a matter within the competence of lay people" (*R. v. Marquard*, 1993, p. 248), and that judges and jurors, equipped with their life experience, have all the tools they need to adequately assess the credibility of witnesses should be cause for alarm.

The Use of False Beliefs and Inappropriate Stereotypes

Adversarial justice systems have long placed considerable emphasis on the demeanor of witnesses to assess their credibility (Blumenthal, 1993; Imwinkelried, 1985; Minzner, 2008; Morrison, Porter, & Fraser, 2007; O'Regan, 2017; Timony, 2000; Wellborn, 1990). However, while the role of nonverbal communication in courtrooms is substantial as in any other face-to-face interactions (Burgoon, Guerrero, & Floyd, 2010; Knapp & Hall, 2010; Levenson, 2008; Remland, 1994; Searcy, Duck, & Blanck, 2005), the misuse of nonverbal communication in criminal, civil, or family trials can have far more serious consequences than in everyday conversations. False beliefs and inappropriate stereotypes about the demeanor of witnesses can adversely influence a trial's outcome. Empirical research has long documented such influence in laboratory court setting experiments (e.g., Bodenhausen, 1988; Bothwell & Jalil, 1992; Pryor & Buchanan, 1984; Rogers, Fox, & Herlihy, 2015; Winkel, & Koppelaar, 1991). However, when it comes to examining the full extent of the influence of false beliefs and inappropriate stereotypes about the demeanor of witnesses in real-life court proceedings, jury trials and bench trials offer different kinds of information. With regard to jury trials, anecdotal evidence offers glimpses of how credibility and deception judgments can be influenced by such false beliefs and inappropriate stereotypes (Heath, 2009). For example, in the US, the post-trial interview of a juror in the case of Ronald Cotton, a man exonerated with the help of DNA evidence after serving more than 10 years in prison for a rape he did not commit (O'Neil, 2001), revealed how his facial expression might have played a part in his conviction:

> He had no change of emotions for eight days. He never changed his facial expression. This was extremely strange to me and, as time went by, I expected to see him react and I never did. And so he seemed more guilty and guiltier and guiltier as time went by. (Loeterman, 1997)

However, in bench trials, when judges have the duty to hand down a written judgment, the influence of false beliefs and inappropriate stereotypes about the demeanor of witnesses can emerge clearly from the written judgment. For example, after reviewing approximately 300 written judgments from courts of Quebec, that is, a Canadian province under the jurisdiction of the Supreme Court of Canada, Denault (2015) concluded that "attention paid to nonverbal behaviors by many decision makers has little or no clear

connection with scientifically validated and recognized knowledge" (p. 126, our translation). In a case of sexual misconduct, for example, a judge concluded that the defendant lacked remorse in part because his "non-verbal language" was supposedly not that of someone who is remorseful:

> Nothing in the evidence, in the pre-sentence report or even in the non-verbal language of the accused both at the trial stage and during sentencing representations suggests any concern whatsoever on the part of the defendant towards the victim in the view of the tragedy she experienced. (*R. c. S. B.*, 2006, pp. 10–11, our translation)

However, while displays of remorse by defendants have been shown to be a decisive factor in sentencing decisions (Bandes, 2014; Haney, Sontag, & Constanzo, 1994; Sundby, 1998), "there is no evidence that facial expression, body language, or other physiological markers exist that can identify feelings of remorse" (Bandes, 2016, p. 17). In other words, the assumption that demeanor is a window to the soul of witnesses to reliably identify remorse is not only baseless, but also highly problematic, more so considering the potential influence of racial and cultural biases (Bandes, 2014, 2016).

In a drug-related case, another judge assumed that the defendant was dishonest because of confusion during his testimony and his "body language." However, it is impossible to know if the judge took into account nonverbal behaviors supported by empirical research rather than false beliefs and inappropriate stereotypes about the demeanor of witnesses:

> Here it appears that the way the defender testifies is such that his version cannot be believed from the outset; his body language as well as his confused explanations have, in some respect, cast doubt on his sincerity. (*R. c. Pinard,* 2014, p. 7, our translation)

Given the absence of clearly diagnostic nonverbal behaviors akin to Pinocchio's nose (DePaulo et al., 2003; Vrij, 2008), the assumption that the "body language" of the defendant reveals dishonesty is at least questionable. Such an assumption is even more doubtful considering the prominence of false beliefs and inappropriate stereotypes among legal practitioners (Denault, 2015, 2017; Strömwall & Granhag, 2003; Vrij, 2008) and their difficulty to differentiate science from unwarranted claims (Fraigman, 2006; Lilienfeld & Landfield, 2008; Moreno, 2003; Tadei, Finnilä, Reite, Antfolk, & Santtila, 2016). In addition, the judges cited above in both the sexual misconduct and drug-related cases referred to the overall demeanor of parties to a dispute. There is no way of knowing what particular nonverbal behaviors influenced them. However, at times, judges refer to specific nonverbal behaviors in written judgments, and the influence of false beliefs and inappropriate stereotypes about the demeanor of witnesses emerge more clearly. For example, in an action for damages caused by vandalism, a judge concluded that the

defendants were honest in part because of their continuous visual contact when testifying:

> The court carefully observed the non-verbal behavior [of the witnesses] during their testimony. These young men responded frankly and spontaneously to questions by looking in the eyes of the president of the tribunal during their testimony. They offered a credible testimony that nothing justifies dismissing. (*Bessette c. Brisson*, 2004, p. 4, our translation)

In an impaired driving action, another judge concluded that the defendant was dishonest while testifying not only because of his nervousness and hesitation, but also because of his gaze aversion:

> Having carefully observed the accused during his testimony and noted his great nervousness, his fleeting glare and his numerous hesitations in cross-examination, the court is convinced that [the defendant] has simply forged his version of the facts according to the evidence disclosed, and that he thereby lied to the court in a shameless manner. (*R. c. Martin*, 2017, p. 27, our translation)

In the above actions for damages and impaired driving, continuous visual contact with the judge was interpreted as a sign of honesty while gaze aversion, nervousness, and hesitation as signs of dishonesty. However, both liars and truth-tellers can be nervous and hesitate. In addition, regarding eye behavior, the interpretation of the judges in the two above actions runs contrary to countless studies of empirical research finding no difference in gaze aversion between truth-tellers and liars (DePaulo et al., 2003; Vrij, 2008). In fact, liars might even use more deliberate eye contact than truth-tellers not only to countermeasure the unwarranted claim that gaze aversion is a cue to deceit, but also to see reactions from their conversational partners and appear credible (Mann et al., 2012, 2013).

Obviously, it is virtually impossible to know for sure the full extent of the influence of false beliefs and inappropriate stereotypes on a trial's outcome nor the accuracy of the credibility assessment of witnesses. However, the demeanor of witnesses prevails as one of the most important factors to assess credibility which, in turn, is the "largest determinant of a deception judgment" (Bond & DePaulo, 2008, p. 487). Therefore, it is safe to say that false beliefs and inappropriate stereotypes about the demeanor of witnesses can have a significant impact on deception judgments of judges and jurors. This impact can be especially salient when there is no evidence proving or disproving the version of witnesses, in what are called "he said-she said" trials (Porter et al., 2003; Seniuk, 2013). Therefore, considering the legion of other extra-legal factors (e.g., sexual orientation, weight, facial appearance, and gender) that can potentially influence the credibility of witnesses, one cannot deny the fact that a trial's outcome is prone to dubious distortions (Baker, Porter, ten Brinke, & Mundy, 2016; Beety, 2013; Dumas & Testé, 2006; Eberhardt,

Davies, Purdie-Vaughns, & Johnson, 2006; Porter & ten Brinke, 2009; Ragaz & Russell, 2010; Seelau & Seelau, 2005; Wilson & Rule, 2015, 2016; Zebrowitz & McDonald, 1991). In fact, "with all else being the same, when one judge convicts where another would have acquitted, the actual determination of guilt or innocence is essentially left to the 'luck of the draw' of judges" (Seniuk, 1992, p. 85).

Limitations to Using Novel Deception Detection Techniques During Trials

When it comes to improving deception detection accuracy of judges or jurors, perhaps one of the first things scholars and legal practitioners might do is turn to novel deception detection techniques that succeeded in improving deception detection accuracy (Hartwig, Granhag, & Luke, 2014; Nahari, Vrij, & Fisher, 2012; Oberlander et al., 2016; Vrij, 2005; Vrij, Fisher, & Blank, 2017; Vrij & Granhag, 2012a; Vrij, Mann, & Fisher, 2006). Given the absence of diagnostic nonverbal behaviors (DePaulo et al., 2003; Vrij, 2008), relying solely on the facial expression of witnesses, for example, would be deemed naïve and misinformed. However, in courtrooms of adversarial justice systems, the implementation of novel deception detection techniques to elicit and enhance cues to deceit raises questions (Masip, 2017; Vrij & Granhag, 2012a, 2012b). For example, in a recent article asking which lie detection tools are ready for the criminal justice system, Vrij and Fisher (2016) list many new approaches such as increasing cognitive load by inviting respondents to do a concurrent task, asking unanticipated questions, using evidence strategically, or requesting verifiable details. However, these new approaches are meant to be used in investigative interview settings but are inappropriate for real-life court proceedings. Characteristics of trials (e.g., the rules of evidence and the judge's and jury's roles) make these new approaches of little to no value for judges and jurors in criminal, civil, and family trials.

For example, the judge's and jury's roles limit the use of techniques that require one to ask specific open-ended and follow-up questions (Vrij, Fisher, & Blank, 2017; Vrij & Granhag, 2012a; Vrij, Mann, & Fisher, 2006) or written declarations (Oberlander et al., 2016; Vrij, 2005). As aforementioned, in courtrooms of adversarial justice systems, parties to a dispute each have to present to a judge or jury their own evidence and arguments in support of their position (Sward, 1989). Judges or jurors "wait passively for what the parties will present, almost never knowing—often not suspecting—what the parties have chosen not to present" (Frankel, 1975, p. 1038). However, judges can sometimes turn from being passive to being active (Eades, 2008; Paciocco, 2010). In Canada, for example, judges can ask questions on occasions so that justice is done, but questions should not raise doubts as to their impartiality:

…it is clear that judges are no longer required to be as passive as they once were; to be what I call sphinx judges. We now not only accept that a judge may intervene in the adversarial debate, but also believe that it is sometimes essential for him to do so for justice in fact to be done. Thus a judge may and sometimes must ask witnesses questions, interrupt them in their testimony and if necessary call them to order. (*Brouillard Also Known As Chatel v. The Queen*, 1985, p. 44)

However, even if judges have "not only the right, but also the duty to put questions to a witness in order to clarify an obscure answer or to resolve possible misunderstanding of any question by a witness, even to remedy an omission of counsel […] in order to bring out or explain relevant matters" (*R. v. Darlyn*, 1946, p. 487; *Brouillard Also Known As Chatel v. The Queen*, 1985, p. 46), they are not investigators (Gerber, 1987; Sward, 1989). Their role is "to hear and determine the issues raised by the parties, not to conduct an investigation or examination on behalf of society at large" (*Jones v. National Coal Board*, 1957, p. 158; *Brouillard Also Known As Chatel v. The Queen*, 1985, p. 44). Therefore, if they take control over the development and the presentation of a case and subject a defendant to their own cross-examination, judges risk having their verdicts overturned by appellate courts (Monaghan, 2015; Paciocco, 2010). Consequently, at the present time, deception detection techniques to elicit and enhance cues to deceit that require one to ask specific open-ended and follow-up questions (Vrij, Fisher, & Blank, 2017; Vrij & Granhag, 2012a; Vrij, Mann, & Fisher, 2006) or written declarations (Oberlander et al., 2016; Vrij, 2005) cannot be used by jurors and are unlikely to be used by judges in bench trials. Similar restraints apply to techniques that require one to withhold (Hartwig, Granhag, & Luke, 2014) or verify information (Nahari, Vrij, & Fisher, 2012).

In criminal proceedings, for example, prosecutors have an obligation of disclosure to defendants. In Canada, the obligation of disclosure includes all information that "can reasonably be used by the accused either in meeting the case for the Crown, advancing a defence or otherwise in making a decision which may affect the conduct of the defence" (*R. v. Egger*, 1993, p. 467). In other words, "if there is a reasonable possibility that the withholding of [the] information will impair the right of the accused to make full answer and defence" (*R. v. Stinchcombe*, 1991, p. 340), the information has to be disclosed, a feature echoing that of the Supreme Court of the US in *Brady v. Maryland* (1963) and *Giglio v. United States* (1972) (Koppell, 2014). With regard to civil and family proceedings, depending on the jurisdiction, parties to a dispute can obtain incriminating evidence through pretrial discovery. Moreover, during trials, parties to the dispute can be required not to disclose new evidence to take others by surprise (Sward, 1989; Tanford, 2009). According to *Quebec's Code of Civil Procedure*, for example, parties to a dispute have an obligation to disclose the evidence they are planning to use:

> A pleading must specify its nature and purpose and state the facts on which it is based and the conclusions sought. It must also state anything which, if not alleged, could take another party by surprise or raise an unexpected debate. The statements it contains must be clear, precise and concise, presented in logical order and numbered consecutively. (2018, section 99)

Thus, as defendants in criminal trials, parties to a dispute in civil and family trials are typically aware in advance of incriminating evidence. Moreover, judges or jurors are not supposed to possess information unknown to parties to the dispute and cannot pause trials to obtain additional information from a third party in order to evaluate the answers of witnesses (Gerber, 1987; Sward, 1989). Consequently, at the present time, because of the rules of evidence, deception detection techniques to elicit and enhance cues to deceit that require one to withhold (Hartwig, Granhag, & Luke, 2014) or verify information (Nahari, Vrij, & Fisher, 2012) cannot be used by jurors and are unlikely to be used by judges in bench trials. Therefore, when it comes to improving the deception detection accuracy of judges or jurors, the rules of evidence and the judge's and jury's roles make various techniques developed in investigative interview settings, such as those listed by Vrij and Fisher (2016), of little to no value. In other words, even if novel deception detection techniques to elicit and enhance cues to deceit succeeded in improving deception detection accuracy in investigative interview settings, scholars and legal practitioners should exercise caution before trying to implement them in criminal, civil, and family trials. Accordingly, it goes without saying that other procedures dismissed for investigative interviews (e.g., examining nonverbal behaviors) (Vrij & Granhag, 2012a, 2012b) should not be hastily dismissed for judges or jurors, more so considering that deception detection in courtrooms where multiple witnesses and lawyers use deception strategies of all kinds has yet to be addressed.

CONCLUDING REMARKS: A CALL TO DECEPTION DETECTION SCHOLARS

While there is no doubt that the complexity of criminal, civil, and family trials in adversarial justice systems might act as a deterrent for scholars, the consequences of overlooking deception detection in courtrooms should act as a driving force. Judges and jurors are often ill-equipped to understand human behavior (Bond & DePaulo, 2008; Denault, 2015; Porter & ten Brinke, 2009; Strömwall & Granhag, 2003; Vrij, 2008). The risk that a perpetrator of a crime may potentially be acquitted while an innocent individual is convicted based on dubious deception judgments made by judges or jurors is substantial (Denault & Dunbar, 2017; Denault & Jupe, 2017). In civil and family trials, while capital punishment is not at stake, undetected deception can result in disastrous financial and human consequences. Considering several million civil and family cases are filed every year throughout the world

(Clark, 1990; Clemenz & Gugler, 2000; Ramseyer & Rasmusen, 2013; Yates, Davis, & Glick, 2001), such consequences should motivate scholars to take action to study credibility assessment in courtrooms more fully.

Furthermore, since research on deception detection in courtrooms remains underexplored, one could argue that scholars should first engage in dialogue with legal practitioners to develop a detailed understanding of criminal, civil, and family trials. As aforementioned, if characteristics of trials (e.g., the rules of evidence and the judge's and jury's roles) are inadequately understood, scholars might provide inappropriate advice to legal practitioners (Denault & Jupe, 2018; Denault, Jupe, Rochat, & Dodier, 2017). Thorough descriptive research on deception in courtrooms should first be conducted followed by laboratory experiments that mimic real-life court proceedings to test deception detection research hypotheses (Denault et al., 2017; Park, Levine, McCornack, Morrison, & Ferrera, 2002; Rozin, 2001; Serota, Levine, & Boster, 2010). Understanding deception in courtrooms before trying to detect it appears to be an important step to take in order to mitigate the drawbacks of creating laboratory experiments that lack ecological validity (Levine, 2018). Such drawbacks could be further mitigated by promoting interdisciplinary research teams, with scholars and legal practitioners, to develop research questions as relevant as possible for criminal, civil, and family trials. For example, many mock crime studies conducted over the years have taught us many lessons about the efficacy of interviewing approaches (e.g., Gödert, Gamer, Rill, & Vossel, 2005; Vrij, Mann, Kristen, & Fisher, 2007). Once characteristics of trials are adequately understood, the same could be done using mock trials to learn effective techniques for detecting deception in courtrooms (e.g., Burnett & Badzinski, 2005; Talwar & Crossman, 2012). While such initial steps will likely require the involvement of a great number of deception detection scholars and take many years to achieve, one should keep in mind that every day that passes is a day judges and jurors likely misuse nonverbal communication to infer whether witnesses are lying or telling the truth. Moreover, when these court decisions become final and enforceable, they can set legal precedents, and other judges or jurors can subsequently be compelled to use false beliefs and inappropriate stereotypes.

The motivation to take action should be all the greater since deception detection techniques to elicit and enhance cues to deceit that are of little to no value for judges or jurors are unlikely to be used by lawyers to improve the deception detection accuracy of judges and jurors in courtrooms of adversarial justice systems:

> The business of the advocate, simply stated, is to win if possible without violating the law. (The phrase "if possible" is meant to modify what precedes it, but the danger of slippage is well known.) His is not the search for truth as such. To put that thought more exactly, the truth and victory are mutually incompatible for some considerable percentage of the attorneys trying cases at any given time. (Frankel, 1975, p. 1037)

In other words, even if "to get at the truth and arrive at the right result […] is the sole objective of the judge, and counsel should never lose sight of that objective in thinking that the end purpose is to win for his side" (Peck, 1954, p. 433), adversarial justice systems encourage "people actively to cover up facts that could lead to a more accurate portrayal of truth" (Sward, 1989, p. 317). The "advocate's prime loyalty is to his client, not to truth as such" (Frankel, 1975, p. 1035). Thus, we hope that this chapter encourages scholars concerned with the search for truth and justice to give serious consideration to the study of deception detection in courtrooms. Given the high stakes of criminal, civil, and family trials, research is needed to advise legal practitioners more adequately.

REFERENCES

Aetna Life Ins. Co. v. Ward, 140 U.S. 76 (1891).

Atkinson, J. M., & Drew, P. (1979). *Order in court: The organization of verbal interaction in judicial settings*. London: Macmillan Press.

Baker, A., Porter, S., Ten Brinke, L., & Mundy, C. (2016). Seeing is believing: Observer perceptions of trait trustworthiness predict perceptions of honesty in high-stakes emotional appeals. *Psychology, Crime, & Law, 22*(9), 817–831.

Bandes, S. A. (2014). Remorse, demeanor, and the consequences of misinterpretation: The limits of law as a window to the soul. *Journal of Law, Religion and State, 3*(2), 170–199.

Bandes, S. A. (2016). Remorse and criminal justice. *Emotion Review, 8*(1), 14–19.

Beety, V. E. (2013). Criminality and corpulence: Weight bias in the courtroom. *Seattle Journal for Social Justice, 11*(2), 523–554.

Bell, E. (2013). An introduction to judicial fact-finding. *Commonwealth Law Bulletin, 39*(3), 519–552.

Bell, V., Villalobos, J. G., & Davis, D. (2014). Attorneys. In T. Levine (Ed.), *Encyclopedia of deception* (pp. 41–45). Los Angeles: Sage.

Bennett, M. W. (2015). Unspringing the witness memory and demeanor trap: What every judge and juror needs to know about cognitive psychology and witness credibility. *American University Law Review, 64*(6), 1331–1376.

Bessette c. Brisson, 2004 CanLII 44897 (QC CQ).

Blumenthal, J. A. (1993). A wipe of the hands, a lick of the lips: The validity of demeanor evidence in assessing witness credibility. *Nebraska Law Review, 72*(4), 1157–1204.

Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology, 55*(5), 726–737.

Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477–492.

Bothwell, R., & Jalil, M. (1992). The credibility of nervous witnesses. *Journal of Social Behavior and Personality, 7,* 581–586.

Brady v. Maryland, 373 U.S. 83 (1963).

Brodsky, S. L., & Pivovarova, E. (2016). The credibility of witnesses. In C. Willis-Esqueda & B. H. Bornstein (Eds.), *The witness stand and Lawrence S. Wrightsman, Jr.* (pp. 41–52). New York: Springer.

Brouillard Also Known As Chatel v. The Queen, [1985] 1 SCR 39, 1985 CanLII 56 (SCC).

Browning, J. (2014). Snitches get stitches: Witness intimidation in the age of Facebook and Twitter. *Pace Law Review, 35*(1), 192–214.

Buel, S. M. (2014). De facto witness tampering. *Berkeley Journal of Gender, Law and Justice, 29*(1), 72–131.

Buller, D. B., & Burgoon, J. K. (1994). Deception: Strategic and nonstrategic communication. In J. A. Daly & J. M. Wiemann (Eds.), *Strategic interpersonal communication* (pp. 191–223). Hillsdale: Erlbaum.

Buller, D., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203–242.

Buller, D. B., Burgoon, J. K., White, C., & Ebesu, A. (1994). Interpersonal deception: VII. Behavioral profiles of falsification, equivocation and concealment. *Journal of Language and Social Psychology, 13*(4), 366–395.

Bulow-Moller, A. M. (1991). Trial evidence: Overt and covert communication in court. *International Journal of Applied Linguistics, 1*(1), 38–60.

Burgoon, J. K., Blair, J. P., & Strom, R. E. (2008). Cognitive biases and nonverbal cue availability in deception detection. *Human Communication Research, 34*(4), 572–599.

Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2010). *Nonverbal communication*. Boston: Pearson.

Burnett, A., & Badzinski, D. M. (2005). Judge nonverbal communication on trial: Do mock trial jurors notice? *Journal of Communication, 55*(2), 209–224.

Clark, D. S. (1990). Civil litigation trends in Europe and Latin America since 1945: The advantage of intracountry comparisons. *Law & Society Review, 24*(2), 549–570.

Clemenz, G., & Gugler, K. (2000). Macroeconomic development and civil litigation. *European Journal of Law and Economics, 9*(3), 215–230.

Connick, E., & Davis, R. C. (1983). Examining the problem of witness intimidation. *Judicature, 66*(9), 439–448.

Canada's Criminal Code, R.S.C. 1985, c. C-46. Retrieved from http://laws-lois.justice.gc.ca/eng/acts/C-46/.

Denault, V. (2015). *Communication non verbale et crédibilité des témoins* [Nonverbal communication and the credibility of witnesses]. Cowansville: Éditions Yvon Blais.

Denault, V. (2017). Le "langage" non verbal des témoins, quand les pseudosciences s'invitent au tribunal [The "body language" of witnesses, when pseudosciences are invited in the courtroom]. *ScriptUM*: *La revue du Colloque VocUM 2015, 2,* 96–118.

Denault, V., & Dunbar, N. (2017). Nonverbal communication in courtrooms: Scientific assessments or modern trials by ordeal? *The Advocates' Quarterly, 47*(3), 280–308.

Denault, V., & Jupe, L. (2017). Deception detection. In B. Baker, R. Minhas, & L. Wilson (Eds.), *Psychology and law factbook 2*. Derby: European Association of Psychology and Law Student Society.

Denault, V., & Jupe, L. (2018). Detecting deceit during trials: Limits in the implementation of lie detection research—A comment on Snook, McCardle, Fahmy and House. *Canadian Criminal Law Review, 23*(1), 97–106.

Denault, V., Jupe, L., Dodier, O., & Rochat, N. (2017). To veil or not to veil, detecting lies in the courtroom: A comment on Leach et al. (2016). *Psychiatry, Psychology and Law, 24*(1), 102–117.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74*(1), 63–79.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979–995.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–112.

Doyon, F. (1999). L'évaluation de la crédibilité des témoins [The credibility assessment of witnesses]. *Revue canadienne de droit pénal, 4*, 331–343.

Dumas, R., & Testé, B. (2006). The influence of criminal facial stereotypes on juridic judgments. *Swiss Journal of Psychology, 65*(4), 237–244.

Eades, D. (2008). Telling and retelling your story in court: Questions, assumptions and intercultural implications. *Current Issues in Criminal Justice, 20*(2), 209–230.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*(5), 383–386.

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, marriage, and politics.* New York: W. W. Norton.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32,* 88–106.

Farmer, C., & Hancock, J. (2014). Perjury. In T. Levine (Ed.), *Encyclopedia of deception* (pp. 753–756). Los Angeles: Sage.

Fawcett, H. E. (2014). Witness, false testimony of. In T. Levine (Ed.), *Encyclopedia of deception* (pp. 937–940). Los Angeles: Sage.

Fortune, W. H., Underwood, R. H., & Imwinkelried, E. J. (1996). *Modern litigation and professional responsibility handbook.* New York: Aspen Publishers.

Fraigman, D. L. (2006). Judges as amateur scientists. *Boston University Law Review, 86*(5), 1207–1226.

Frankel, M. E. (1975). The search for truth: An umpireal view. *University of Pennsylvania Law Review, 123*(5), 1031–1059.

Friedland, S. I. (1989). On common sense and the evaluation of witness credibility. *Case Western Reserve Law Review, 40*(1), 165–226.

Galasinski, D. (2000). *The language of deception: A discourse analytical study.* Thousand Oaks: Sage.

Gerber, R. J. (1987). Victory vs. truth: The adversary system and its ethic. *Arizona State Law Journal, 19*(1), 3–26.

Gödert, H. W., Gamer, M., Rill, H. G., & Vossel, G. (2005). Statement validity assessment: Inter-rater reliability of criteria-based content analysis in the mock-crime paradigm. *Legal and Criminological Psychology, 10*(2), 225–245.

Giglio v. United States, 405 U.S. 150 (1972).

Granhag, P. A., & Strömwall, L. A. (2004). *The detection of deception in forensic contexts.* Cambridge: Cambridge University Press.

Green, S. P. (2001). Lying, misleading, and falsely denying: How moral concepts inform the law of perjury, fraud, and false statements. *Hasting Law Journal, 53*(1), 157–212.

Griffin, L. K. (2013). Narrative, truth, and trial. *Georgetown Law Journal, 101*(2), 281–336.

Haney, C., Sontag, L., & Constanzo, S. (1994). Deciding to take a life: Capital juries, sentencing instructions, and the jurisprudence of death. *Journal of Social Issues, 50*(2), 149–176.

Hartwig, M., Granhag, P. A., & Luke, T. (2014). Strategic use of evidence during investigative interviews: The state of the science. In D. C. Raskin, C. R. Honts, & J. C. Kircher (Eds.), *Credibility assessment: Scientific research and applications* (pp. 1–36). Oxford: Academic Press.

Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review, 19*(4), 307–342.

Heath, W. P. (2009). Arresting and convicting the innocent: The potential role of an "inappropriate" emotional display in the accused. *Behavioral Sciences and the Law, 27*(3), 313–332.

Hutchins, R. M. (2014). You can't handle the truth! Trial juries and credibility. *Seton Hall Law Review, 44*(2), 505–556.

Imwinkelried, E. J. (1985). Demeanor impeachment: Law and tactics. *American Journal of Trial Advocacy, 9*(2), 183–236.

Jones v. National Coal Board, [1957]. 2 All E.R. 155 (C.A.).

Kane, J. L. (2007). Judging Credibility. *Litigation, 33*(3), 31–37.

Knapp, M. L., & Hall, J. A. (2010). *Nonverbal communication in human interaction.* Boston: Wadsworth.

Köhnken, G. (1989). Behavioral correlates of statement credibility: Theories, paradigms, and results. In H. Wegener, F. Lösel, & J. Haisch (Eds.), *Criminal behavior and the justice system: Psychological perspectives* (pp. 271–289). New York: Springer.

Koppell, S. (2014). An argument against increasing prosecutors' disclosure requirements beyond Brady. *Georgetown Journal of Legal Ethics, 27*(3), 643–654.

Lacy, J. W., & Stark, C. E. L. (2013). The neuroscience of memory: Implications for the courtroom. *Nature Reviews Neuroscience, 14,* 649–658.

Levenson, L. L. (2008). Courtroom demeanor: The theater of the courtroom. *Minnesota Law Review, 92*(3), 573–633.

Levine, E. E., & Schweitzer, M. E. (2014). Are liars ethical? On the tension between benevolence and honesty. *Journal of Experimental Social Psychology, 53,* 107–117.

Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes, 125,* 88–106.

Levine, T. R. (2010). A few transparent liars. *Communication Yearbook, 34*(1), 40–61.

Levine, T. R. (2018). Ecological validity and deception detection research design. *Communication Methods and Measures, 12*(1), 45–54. https://doi.org/10.1080/19312458.2017.1411471.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., & Lee, J. H. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research, 37*(3), 377–403.

Lilienfeld, S. O., & Landfield, K. (2008). Science and pseudoscience in law enforcement: A user-friendly primer. *Criminal Justice and Behavior, 35*(10), 1215–1230.

Loeterman, B. (1997, February 25). What Jennifer saw. *Frontline*. Retrieved from https://www.pbs.org/wgbh/pages/frontline/shows/dna/etc/script.html.

Loevy, J. (2006). How to convince the court that the cops are lying. *Litigation, 32*(2), 33–39.

Mann, S., Ewens, S., Shaw, D., Vrij, A., Leal, S., & Hillman, J. (2013). Lying eyes: Why Liars seek deliberate eye contact. *Psychiatry, Psychology and Law, 20*(3), 452–461.

Mann, S., Vrij, A., Leal, S., Granhag, P. A., Warmelink, L., & Forrester, D. (2012). Windows to the soul? Deliberate eye contact as a cue to deceit. *Journal of Nonverbal Behavior, 36*(3), 205–215.

Masip, J. (2017). Deception detection: State of the art and future prospects. *Psicothema, 29*(2), 149–159.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs, 59*(1), 1–16.

Minzner, M. (2008). Detecting lies using demeanor, bias and context. *Cardozo Law Review, 29*(6), 2557–2582.

Monaghan, N. (2015). *Law of evidence*. Cambridge: Cambridge University Press.

Moreno, J. A. (2003). Einstein on the bench: Exposing what judges do not know about science and using child abuse cases to improve how courts evaluate scientific evidence. *Ohio State Law Journal, 64*(2), 351–584.

Morrison, B. R., & Comeau, W. (2002). Judging credibility of witnesses. *The Advocates Quarterly, 25*(4), 411–440.

Morrison, B. R., Porter, L. L., & Fraser, I. H. (2007). The role of demeanour in assessing the credibility of witnesses. *The Advocates Quarterly, 33*(1), 170–192.

Nahari, G., Vrij, A., & Fisher, R. P. (2012). Exploiting liars verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology, 19*(2), 227–239.

Norris, J., & Edwardh, M. (1995). Myths, hidden facts and common sense: Expert opinion evidence and the assessment of credibility. *Criminal Law Quarterly, 38*(1), 73–103.

Oberlander, V., Naefgen, C., Koppehele-Gossel, J., Quinten, L., Banse, R., & Schmidt, A. F. (2016). Validity of content-based techniques to distinguish true and fabricated statements: A meta-analysis. *Law and Human Behavior, 40*(4), 440–457.

O'Neill, H. (2001, March 4). The perfect witness. *The Washington Post*. Retrieved from http://www.washingtonpost.com/archive/lifestyle/2001/03/04/the-perfect-witness/a7fa0461-c15c-4237-86db-52ab5069fbea.

O'Regan, D. (2017). Eying the body: The impact of classical rules for demeanor credibility, bias, and the need to blind legal decision makers. *Pace Law Review, 37*(2), 379–454.

P. (D.) v. S. (C.), [1993] 4 SCR 141, 1993 CanLII 35 (SCC).

Paciocco, D. M. (2010). Understanding the accusatorial system. *Canadian Criminal Law Review, 14*(3), 307–325.

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrerra, M. (2002). How people really detect lies. *Communication Monographs, 69*(2), 144–157.

Peck, D. W. (1954). The complement of court and counsel. *Record of the Association of the Bar of the City of New York, 9*(6), 272–286.

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34,* 243–281.

Porter, S., & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology, 14*(1), 119–134.

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime, & Law, 16,* 477–491.

Porter, S., Campbell, M. A., Birt, A. R., & Woodworth, M. T. (2003). "He said, she said": A psychological perspective on historical memory evidence in the courtroom. *Canadian Psychology, 44*(3), 190–206.

Pryor, B., & Buchanan, R. W. (1984). The effects of a defendant's demeanor on juror perceptions of credibility and guilt. *Journal of Communication, 34*(3), 92–99.

Quebec's Code of Civil Procedure. 2018. c. C-25.01. Retrieved from http://legisque-bec.gouv.qc.ca/en/showdoc/cs/C-25.01.

R. v. Béland, [1987] 2 SCR 398, 1987 CanLII 27 (SCC).

R. v. Brooks, [2000] 1 SCR 237, 2000 SCC 11 (CanLII).

R. v. Darlyn, [1946] CanLII 248 (BC CA).

R. v. Egger, [1993] 2 SCR 451, 1993 CanLII 98 (SCC).

R. v. François, [1994] 2 SCR 827, 1994 CanLII 52 (SCC).

R. v. Gagnon, [2006] 1 SCR 621, 2006 SCC 17 (CanLII).

R. v. Handy, [2002] 2 SCR 908, 2002 SCC 56 (CanLII).

R. v. Marquard, [1993] 4 SCR 223, 1993 CanLII 37 (SCC).

R. c. Martin, 2017 QCCS 193 (CanLII).

R. c. Pinard, [2014] QCCQ 5630 (CanLII).

R. c. S. B., [2006] QCCQ 12796 (CanLII).

R. v. S. (R.D.), [1997] 3 SCR 484, 1997 CanLII 324 (SCC).

R. v. Stinchcombe, [1991] 3 SCR 326, 1991 SCC 45 (CanLII).

R. v. W. (R.), [1992] 2 SCR 122, 1992 CanLII 56 (SCC).

Ragaz, L. L., & Russell, B. (2010). Sex, sexual orientation, and sexism: What influence do these factors have on verdicts in a crime-of-passion case? *Journal of Social Psychology, 150*(4), 341–360.

Ramseyer, J. M., & Rasmusen, E. B. (2013). *Comparative litigation rates* (Harvard John M. Olin Discussion Paper Series, No. 681, Nov. 2010). Retrieved from http://www.law.harvard.edu/programs/olin_center/papers/pdf/Ramseyer_681.pdf.

Remland, M. S. (1994). The importance of nonverbal communication in the courtroom. *The New Jersey Journal of Communication, 2*(2), 124–144.

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. In B. Cronin (Ed.), *Annual review of information science and technology* (pp. 307–364). Medford: Information Today.

Rogers, H., Fox, S., & Herlihy, J. (2015). The importance of looking credible: The impact of the behavioural sequelae of post-traumatic stress disorder on the credibility of asylum-seekers. *Psychology, Crime, & Law, 21*(2), 139–155.

Rogers, T., Zeckhauser, R. J., Gino, F., Schweitzer, M. E., & Norton, M. I. (2017). Artful paltering: The risks and rewards of using truthful statements to mislead others. *Journal of Personality and Social Psychology, 112*(3), 456–473.

Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review, 5*(1), 2–14.

Schwelb, F. E. (1989). Lying in court. *Litigation, 15*(2), 3–54.

Searcy, M., Duck, S., & Blanck, P. (2005). Nonverbal communication in the courtroom and the "appearance" of justice. In R. E. Riggio & R. S. Feldman (Eds.), *Applications of nonverbal communication* (pp. 41–62). Mahwah: Lawrence Erlbaum.

Seelau, S. M., & Seelau, E. P. (2005). Gender-role stereotypes and perceptions of heterosexual, gay and lesbian domestic violence. *Journal of Family Violence, 20*(6), 363–371.

Seniuk, G. T. G. (1992). Judicial fact-finding and a theory of credit. *Saskatchewan Law Review, 56*(1), 79–112.

Seniuk, G. T. G. (2013). Credibility assessment, common law trials and fuzzy logic. In B. S. Copper, D. Griesel, & M. Ternes (Eds.), *Applied issues in in investigative interviewing, eyewitness memory, and credibility assessment* (pp. 19–30). New York: Springer.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36*(1), 2–25.

Simon-Kerr, J. (2015). Systematic lying. *William & Mary Law Review, 56*(6), 2175–2234.

Slobogin, C. (1996). Testilying: Police perjury and what to do about it. *University of Colorado Law Review, 67*(4), 1037–1060.

Strier, F. (1994). Making jury trials more truthful. *University of California, Davis Law Review, 30*(1), 95–182.

Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime, & Law, 9*(1), 19–36.

Summers, R. S. (1999). Formal legal truth and substantive truth in judicial fact-finding—Their justified divergence in some particular cases. *Law and Phylosophy, 18*(5), 497–511.

Sundby, S. (1998). The capital jury and absolution: The intersection of trial strategy, remorse, and the death penalty. *Cornell Law Review, 83*(4), 1557–1598.

Sward, E. E. (1989). Values, ideology, and the evolution of the adversary system. *Indiana Law Journal, 64*(2), 301–355.

Tadei, A., Finnilä, K., Reite, A., Antfolk, J., & Santtila, P. (2016). Judges' capacity to evaluate psychological and psychiatric expert testimony. *Nordic Psychology, 68*(3), 204–217.

Talwar, V., & Crossman, A. M. (2012). Children's lies and their detection: Implications for child witness testimony. *Developmental Review, 32*(4), 337–359.

Tanford, J. A. (2009). *The trial process: Law, tactics and ethics.* New York: Matthew Bender & Co.

ten Brinke, L., & Porter, S. (2013). Discovering deceit: Applying laboratory and field research in the search for truthful and deceptive behaviour. In B. S. Cooper, D. Griesel, & M. Ternes (Eds.), *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment* (pp. 221–237). New York: Springer.

Timony, J. P. (2000). Demeanor credibility. *Catholic University Law Review, 49*(4), 903–944.

Troville, P. V. (1939). History of lie detection. *Journal of Criminal Law and Criminology, 29*(6), 848–881.

United States v. Scheffer, 523 U.S. 303 (1998).

Van Swol, L. M., & Braun, M. (2014). Communicating deception: Differences in language use, justifications, and questions for lies, omissions, and truths. *Group Decision and Negotiation, 23*(6), 1343–1367.

Vetrovec v. The Queen, [1982] 1 SCR 811, 1982 CanLII 20 (SCC).

Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law, 11*(1), 3–41.

Vrij, A. (2007). Deception: A social lubricant and a selfish act. In K. Fiedler (Ed.), *Frontiers of social psychology: Social communication* (pp. 309–342). New York: Psychology Press.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities.* Chichester: Wiley.

Vrij, A., & Fisher, R. P. (2016). Which lie detection tools are ready for use in the criminal justice system? *Journal of Applied Research in Memory and Cognition, 5*(3), 302–307.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22*(1), 1–21.

Vrij, A., & Granhag, P. A. (2012a). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*(2), 110–117.

Vrij, A., & Granhag, P. A. (2012b). The sound of critics: New tunes, old tunes, and resistance to play. *Journal of Applied Research in Memory and Cognition, 1*(2), 139–143.

Vrij, A., Mann, S., & Fisher, R. (2006). Information-gathering vs accusatory interview style: Individual Differences in respondents' experiences. *Personality and Individual Differences, 41*(4), 589–599.

Vrij, A., Mann, S., Kristen, S., & Fisher, R. P. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*(5), 499–518.

Wellborn, O. G. (1990). Demeanor. *Cornell Law Review, 76*(5), 1075–1105.

White v. The King, [1947] SCR 268, 1947 CanLII 1 (SCC).

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science, 26*(8), 1325–1331.

Wilson, J. P., & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes. *Social Psychological and Personality Science, 7*(4), 331–338.

Wilthermuth, S. S., Newman, D. T., & Raj, M. (2015). The consequences of dishonesty. *Current Opinion in Psychology, 6,* 20–24.

Winkel, F. W., & Koppelaar, L. (1991). Rape victims' style of self-presentation and secondary victimization by the environment. *Journal of Interpersonal Violence, 6*(1), 29–40.

Wise, T. A. (1845). *Commentary on the Hindu system of medicine.* Calcutta: Thacker and Co.

Yates, J., Davis, B. C., & Glick, H. R. (2001). The politics of torts: Explaining litigation rates in the American states. *State Politics & Policy Quaterly, 1*(2), 127–143.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' babyfaceness and attractiveness on adjudications in small claims courts. *Law and Human Behavior, 15*(6), 603–623.

How Supreme Court Deceptions Inflate Presidential Power

Louis Fisher

The Framers understood that public officials will make errors. That is why we have separation of powers between three branches of the national government, reinforced by a system of checks and balances. What would have surprised the Framers is to have one of the branches make a major error to upset that balance of powers, have the error clearly identified from the start, and yet persist decade after decade until finally corrected 79 years later. That is the story of judicial errors in the 1936 Supreme Court decision of *United States v. Curtiss-Wright*. By misreading a speech by John Marshall when he served in the House of Representatives in 1800, the court attributed to the President "plenary and exclusive" powers over external affairs, a position that can be easily rejected simply by reading the text of Articles I and II of the Constitution. When the court in 2015 corrected the *Curtiss-Wright* error, it decided to create a new model of presidential power that recreates the notion of plenary and exclusive powers over external affairs. The details of that story are set forth in this article.

From their close study of history, the Framers understood that the decision to act in external affairs could not be left solely to the President. They consciously and deliberately rejected the British model that placed foreign affairs wholly in the hands of the executive authority. In 1690, John Locke spoke of three branches of government: legislative, executive, and "federative." By the latter he meant "the power of war and peace, leagues and alliances, and all the

L. Fisher (✉)
Scholar in Residence at The Constitution Project, Washington, DC, USA
e-mail: lfisher11@verizon.net

transactions with all persons and communities without the commonwealth" (Locke, p. 191, § 146). For Locke, the federative power, what we today refer to as foreign affairs, was "almost always united" with the executive. Any separation of the executive and federative powers would invite "disorder and ruin" (Locke, pp. 191–192, §§ 147–148).

A second British model appears in the *Commentaries* (1765), written by Sir William Blackstone. He regarded the king's prerogative as "those rights and capacities the king enjoys alone" (Blackstone, 1765, v. 2, p. 232). Those powers included the right to send and receive ambassadors and of "making war or peace" (p. 233). The king could make "a treaty with a foreign state, which shall irrevocably bind the nation" (p. 244). He could issue letters of marque and reprisal (authorizing private citizens to undertake military actions) and possessed "the sole power of raising and regulating fleets and armies" (pp. 250, 254).

During debates at the Philadelphia convention, the Framers transferred to Congress many of Locke's federative powers and Blackstone's prerogatives. The power to take the country to war was not left solely to the President. On June 1, 1787, James Wilson expressed support for a single executive but "did not consider the Prerogatives of the British Monarch as a proper guide in defining the Executive powers" (Farrand, 1966, v. 1, p. 65). Some of these prerogatives "were of a Legislative nature. Among others that of war" (Farrand, 1966, v. 1, pp. 65–66). Edmund Randolph expressed concern about executive power, calling it "the foetus of monarchy" and rejected the British model "as our prototype" (p. 66). Many of the executive powers that Blackstone placed in the king are vested expressly in Congress, including issuing letters of marque and reprisal and raising and regulating fleets and armies. Under the US Constitution, two-thirds of the Senate must agree to treaties.

To the Framers, war invites executive aggrandizement. In Federalist No. 4, an essay by John Jay issued a blunt warning about executive initiatives in external affairs: "It is too true, however disgraceful it may be to human nature, that nations in general will make war whenever they have a prospect of getting any thing by it; nay, absolute monarchs will often make war when their nations are to get nothing by it, but for purposes and objects merely personal, such as a thirst for military glory, revenge for personal affronts, ambition, or private compacts to aggrandize or support their particular families or partisans." Those and other motives, "which affect only the mind of the sovereign, often lead him to engage in wars not sanctified by justice or the voice and interests of his people" (Wright, 2002, p. 101).

From World War II to the present time, the damage to constitutional government from presidential errors, misjudgments, and deception has been heavy. Peter Shane (2009) has pointed out that

> time and time again, it has become evident that Presidents, left relatively unchecked by dialogue with and accountability to the other two branches, behave disastrously. The new unilateral presidency is thus not appealing either

as constitutional interpretation or as good institutional design. To put the point another way, the Framers got this right. (p. 5)

In another contemporary study, Harold Bruff (2015) offers his judgment on how Presidents interpret their constitutional powers: "Even in ordinary times, our system has recently become similar enough to a permanent constitutional dictatorship to give deep pause" (p. 465).

It is now time to explore how the Supreme Court has analyzed the roles of the elected branches in external affairs. For well over a century, there was no bias that favored independent presidential power. The constitutional authority of Congress was fully understood by the judiciary. At no time did courts attempt to assign to the President an independent and exclusive role. The interpretation of external affairs changed fundamentally in 1936 with errors and misconceptions committed by the Supreme Court in its *Curtiss-Wright* decision.

Initial Judicial Orientation

For more than a century, the Supreme Court interpreted constitutional disputes between the elected branches without favoring presidential power over Congress. The first military initiative was the Quasi-War against France in 1798, given that name because actions were on sea not on land. Congress debated the prospect of war by enacting a number of bills to put the country on a war footing. President John Adams recommended to Congress "effectual measures of defense" (Richardson, 1897–1925, v. 1, p. 226). Congress debated those proposals and enacted several dozen bills to support military action against France (1 Stat. 547–611).

Several decisions by the Supreme Court on the Quasi-War helped clarify the scope of congressional authority over war and the deployment of military force. In 1800 and 1801, the court recognized that Congress could authorize hostilities in two ways: either by a formal declaration of war or by passing legislation to support an undeclared war. As explained in the first decision issued in 1800, military conflicts could be "limited," "partial," and "imperfect" without requiring a formal declaration of war. Justice Samuel Chase noted: "Congress is empowered to declare a general war, or congress may wage a limited war; limited in place, in objects, and in time" (*Bas v. Tingy*, 1800, p. 43). With the Quasi-War, Congress authorized "hostilities on the high seas" and did not provide authority "to commit hostilities on land" (p. 43). In the second case, Chief Justice John Marshall wrote for the court: "The whole powers of war being, by the constitution of the United States, vested in congress, the acts of that body can alone be resorted to as our guides in this inquiry" (*Talbot v. Seeman*, 1801, p. 28). Certainly, there is no hint in those cases of "plenary and exclusive" power of the President in external affairs.

Congressional authority is underscored by a decision issued by the Supreme Court in 1804. In passing legislation for the Quasi-War against

France, Congress authorized the President to seize vessels sailing *to* French ports. However, President Adams issued an order directing American ships to capture vessels sailing *to or from* French ports. Here was a direct collision between the two elected branches. Which position would prevail in court? In a unanimous opinion, Chief Justice Marshall held that in this conflict between statutory policy and a presidential order, the statute prevailed. The proclamation by Adams could not "change the nature of the transaction, or legalize an act which, without those instructions, would have been a plain trespass" (*Little v. Barreme*, 1804, p. 179).

This decision is of special interest because it was written by John Marshall four years after he gave his sole-organ speech as a member of the House of Representatives. In the *Curtis-Wright* decision in 1936, the Supreme Court attempted to interpret his speech as conferring upon the President plenary and exclusive power over external affairs. That was not Marshall's position as a member of Congress or as Chief Justice.

In a significant case interpreting the authority of Congress to restrict presidential action abroad, a circuit court in 1806 reviewed the indictment of Colonel William S. Smith for engaging in military action against Spain. He claimed that his initiative "was begun, prepared, and set on foot with the knowledge and approbation of the executive department of our government." Could a President or his advisers somehow authorize military adventures that violated congressional policy set forth in the Neutrality Act of 1794? In rejecting that argument, the court described the statute as "declaratory of the law of nations; and besides, every species of private and unauthorized hostilities is inconsistent with the principles of the social compact, and the very nature, scope, and end of civil government" (*United States v. Smith*, 1806, p. 1229).

As to Smith's claim that somehow he had the support of the executive branch, the court rejected the notion that the Neutrality Act allowed the President or executive officials to waive statutory policy: "If a private individual, even with the knowledge and approbation of this high and preeminent officer of our government [the President], should set on foot such a military expedition, how can he expect to be exonerated from the obligation of the law?" The President "cannot control the statute, nor dispense with its execution, and still less can he authorize a person to do what the law forbids" (*United States v. Smith*, 1806, p. 1230). Here is very clear language repudiating any notion of the President as sole organ in external affairs. Other courts during this period issued similar rulings until the Supreme Court in *Curtis-Wright* spoke about "plenary and exclusive" power for the President in foreign affairs.

In 1860, a federal court offered support for independent actions by the President to use military force abroad to protect American lives and property. An American ship in the Nicaraguan port of Greytown ordered local authorities to make amends for an alleged affront to an American diplomat.

Unsatisfied by the response, the commander of the American *Cyane* ordered the town bombarded from nine in the morning to midafternoon. American forces went ashore to destroy whatever remained of the town. A federal circuit court described the President as "the only legitimate organ of the general government, to open and carry on correspondence or negotiation with foreign nations, in matters concerning the interests of the country or its citizens." The court concluded that citizens abroad must look to the President "for protection of person and of property" (*Durand v. Hollins*, 1860, p. 112).

In 1868, Congress passed legislation to promote a more balanced policy by requiring the President to demand from a foreign government the reason for depriving any American citizen of liberty. If it appeared wrongful and a violation of the citizens' rights, the President was required to demand the citizen's release. If the foreign government delayed or refused, the President could use such means "not amounting to acts of war" that he thought necessary and proper to obtain the citizen's release (15 Stat. 223). Legislation in 1989 inserted "and not otherwise prohibited by law" after "acts of war" (103 Stat. 1900, sec. 9).

Throughout the 1800s and into the early 1900s, the Supreme Court issued a number of immigration cases to determine limits on state and local government treatment of aliens. The pattern was to defer to the constitutional authority of Congress, relying on the express power of Congress in Article I, Section 8, to "regulate Commerce with foreign Nations, and among the several States." Justices did not recognize any kind of implied, inherent, or exclusive power of the President over external affairs, as would be advanced by the Supreme Court in *Curtiss-Wright* (Fisher, 2017, pp. 47–49).

THE SOLE-ORGAN DOCTRINE

The Supreme Court has a record of placing gratuitous and extraneous material in its decisions. Referred to as "dicta," these statements do not embody the determination of a court and are not binding in subsequent cases. This type of dicta can be erroneous. Nevertheless, litigants, scholars, lower courts, and the Supreme Court rely on dicta as though it is central to the actual holding. To Justice Benjamin Cardozo, it was "a good deal of mystery how judges, of all persons in the world, should put their faith in dicta." There was constant need, he emphasized, "as every law student knows, to separate the accidental and the non-essential from the essential and inherent" (Cardozo, 1921, pp. 29–30). That understanding is regularly ignored by both lower courts and the Supreme Court.

The risk of judicial errors and misconceptions about historical precedents should be well understood. In an article published in 1945, Justice Robert Jackson remarked that judges "often are not thorough or objective

historians" (Jackson, 1945, p. 6). In 1965, Alfred Kelly offered his evaluation of the court's role as constitutional historian: "if not a naked king, no better than a very ragged one. From a professional point of view, most, if not all, of its recent historical essays are very poor indeed." All too often Justices "reach conclusions that are plainly erroneous" (Kelly, 1965, p. 155).

The issue before the Supreme Court in *Curtiss-Wright* concerned legislation passed by Congress in 1934 authorizing the President to prohibit the sale of arms in the Chaco region of South America whenever he found it "may contribute to the reestablishment of peace" between belligerents (48 Stat. 811, ch. 365). In enforcing the embargo, President Franklin D. Roosevelt made no claim of independent or exclusive power over external affairs. His proclamation prohibiting the sale of arms and munitions to the Chaco rested entirely on authority granted him by Congress: "Now, therefore, I, Franklin D. Roosevelt, President of the United States, acting under and by virtue of the authority conferred in me by the said joint resolution of Congress …" (48 Stat. 1745).

In upholding the delegation, Justice George Sutherland chose to go beyond that basic issue put to the court. He decided to announce a broad grant of independent power to the President over external affairs, relying on misconceptions and errors that would shape constitutional law for decades to come. Edward Keynes (1982) described Sutherland's decision "as celebrated and widely accepted," but cautioned that "Sutherland's theory suffers from several empirical and theoretical limitations" (pp. 84–85). A core mistake was the decision of Justice Sutherland to rely on a speech given by John Marshall in 1800 when he served in the House of Representatives.

In 1800, President John Adams ran for reelection. His opponent was Thomas Jefferson. In the House, Jeffersonians urged that Adams be either impeached or censured for turning over to Great Britain an individual charged with murder. Because the case was already pending in an American court, some lawmakers wanted to punish Adams for encroaching on the judiciary and violating the doctrine of separation of powers. According to a House resolution, turning the individual over to the British marked "a dangerous interference of the Executive with Judicial decisions" (10 Annals of Cong. 533). Some lawmakers "had no doubt of the competency of the House either to impeach, to censure, or to approbate the conduct of the Executive" (10 Annals of Cong. 553).

There was confusion about the nationality of the individual turned over to England. The House resolution rebuking Adams began with these words: "it appears to this House that a person, calling himself Jonathan Robbins, and claiming to be a citizen of the United States," was committed to trial in the US "for the alleged crime of piracy and murder, committed on the high seas, on board the British frigate Hermione" (10 Annals of Cong. 532). Notice the words "it appears," "calling himself," and "claiming to be." Robbins told the government he was from Danbury, Connecticut, but citizens living there

certified they had never known an inhabitant of the town "by the name of Jonathan or Nathan Robbins, and that there has not been nor now is any family known by the name of Robbins within the limits of said town" (10 Annals of Cong. 517). Secretary of State Timothy Pickering concluded that Robbins was using an assumed name and was actually Thomas Nash, a native Irishman (10 Annals of Cong. 515). US District Judge Thomas Bee of South Carolina, asked to turn the prisoner over to the British, agreed that the individual was Thomas Nash (10 Annals of Cong. 515).

In a lengthy floor speech, Marshall said that if President Adams had directed Judge Bee to decide "for or against his own jurisdiction, to condemn or acquit the prisoner, this would have been a dangerous interference with judicial decisions, and ought to have been resisted" (10 Annals of Cong. 615–616). But there was no such interference. National policy for external affairs had been made jointly by the President and the Senate through the treaty process. As one reads the speech, it is clear that Adams had turned over to England Thomas Nash, acting under Article 27 of the Jay Treaty, which in 1794 specifically authorized the President to extradite to England British citizens charged either with murder or forgery (8 Stat. 129).

In the course of defending President Adams, Marshall referred to the President as "the sole organ of the nation in its external relations" (10 Annals of Cong. 614). To Sutherland and the Justices who joined the majority opinion in *Curtis-Wright*, that language gave the President plenary and exclusive authority over foreign affairs. Obviously the Justices read a single sentence from Marshall about the President as "sole organ" but did not read the entire speech to understand the context of those two words. They attributed to Marshall a policy he did not embrace. The court attributed to the President a source of power in foreign affairs that did not depend on authority delegated by Congress, but was somehow implied in the Constitution:

> It is important to bear in mind that we are here dealing not alone with an authority vested in the President by an assertion of legislative power, but with such an authority plus the very delicate, plenary and exclusive power of the President as the sole organ of the federal government in the field of international relations—a power which does not require as a basis for its exercise an act of Congress, but which, of course, like every other governmental power, must be exercised in subordination to the applicable provisions of the Constitution. It is quite apparent that if, in the maintenance of our international relations, embarrassment—perhaps serious embarrassment—is to be avoided and success for our aims achieved, congressional legislation which is to be made effective through negotiation and inquiry into the international field must often accord to the President a degree of discretion and freedom from statutory restriction which would not be admissible were domestic affairs alone involved. (*United States v. Curtiss-Wright Corp.*, 1936, pp. 319–320)

John Marshall made no such argument. In defending President Adams, he relied solely on authority granted by the Jay Treaty. Moreover, the elementary

step of reading Articles I and II of the Constitution underscores that much of foreign affairs is shared between the legislative and executive branches. Marshall did not promote some kind of plenary, exclusive, or independent presidential power over external affairs. There was no effort by Adams to make foreign policy unilaterally. He was not the sole organ in formulating the treaty. He was the sole organ in *implementing* it, which was his express Article II constitutional duty to "take Care that the Laws be faithfully executed," whether expressed by statute or by treaty.

Justice Sutherland made two other errors in *Curtiss-Wright*. He claimed that the President "*makes* treaties with the advice and consent of the Senate; but he alone negotiates. In the field of negotiation the Senate cannot intrude; and Congress itself is powerless to invade it" (p. 319, emphasis in original). There is no basis for that understanding. If one wants a good source to discredit Sutherland's statement, it would be his book published in 1919, reflecting his twelve years as a United States Senator from Utah. In that book, Sutherland describes how Senators are regularly involved in treaty negotiation because it helps Presidents build political support for a treaty (Sutherland, 1919, pp. 122–124). Presidents invite not only Senators to participate in treaty negotiation but also members of the House when authorization and appropriation bills are needed to implement treaties (Fisher, 1989).

A third Sutherland error in *Curtiss-Wright* is his claim that after America's war of independence, the powers of external sovereignty passed from the Crown "not to the colonies severally, but to the colonies in their collective and corporate capacity as the United States of America" (p. 316). That is false. After declaring independence from England, several of the colonies acted in a sovereign capacity by entering into treaties. The treaty with Great Britain acknowledged that reality by referring to New Hampshire, Massachusetts Bay, Rhode Island, and the other colonies "to be free, sovereign and independent States" (8 Stat. 55). This historical error by Sutherland was closely analyzed by Julius Goebel, Jr. in an article published in the *Columbia Law Review* in 1938 (Goebel, 1938).

CRITIQUES OF *CURTISS-WRIGHT*

Scholars who closely studied *Curtiss-Wright* have thoroughly discredited the decision for its careless and false mischaracterization of Marshall's sole-organ speech, the misconception about the treaty negotiation process, and misunderstanding the shift of sovereign authority from Great Britain to the United States. An article by David Levitan in 1946 concluded that the "record of events leaves no doubt that treaty-making power was exercised by the States" and that Sutherland's theory that the power of external sovereignty passing from the Crown not to individual colonies but to the colonies in their "collective and corporate capacity as the United States of America" did not "harmonize with the facts. It simply was not so" (p. 489).

In a 1988 article, Michael Glennon referred to the "extravagant scheme concocted by Justice George Sutherland, first unveiled in his earlier writings and later, in 1936, transposed into a Supreme Court opinion, and unleashed upon the nation in *United States v. Curtiss-Wright Export Corp.*" (p. 11). As to Sutherland's sole-organ doctrine and the assertion that the President possesses "plenary powers," Glennon described that as "the sheerest of *dicta*" (p. 12). He regarded Sutherland's opinion as "a muddled law review article wedged with considerable difficulty between the pages of United States Reports" (p. 13).

In a 1988 article, David Gray Adler points out that after deciding that the delegation by Congress to President Roosevelt was not unduly broad, Justice Sutherland strayed from the constitutional issue to select "some ill-considered dicta" that left an unhappy legacy (Adler, 1988, 1, p. 30). As for the sole-organ speech, Adler explains that John Marshall was merely defending the decision of President Adams to act under the Jay Treaty by surrendering to England a British deserter charged with murder. To Adler, "the sole organ doctrine is simply so much fanciful rhetoric" (p. 34).

Notwithstanding these scholarly critiques, an article in 1996 by Anthony Simones points out that "judges have utilized *Curtiss-Wright* to sanction a broad range of presidential powers" (Simones, 1996, p. 411). After reading the repudiations of Sutherland's opinion, Simones expected the decision to be "tossed into the dust bin of constitutional jurisprudence," along with *Dred Scott v. Sandford* and *Plessy v. Ferguson.* Instead, most judges "don't seem to care about the historical basis of Justice Sutherland's theory and don't recall the specific facts of the case" (p. 415). Many other scholarly works have analyzed the weaknesses of *Curtiss-Wright* (Fisher, 2016, pp. 186–199).

From 1936 forward, the White House, the Justice Department, the State Department, and many other executive branch agencies depended heavily on dicta in *Curtiss-Wright* to expand presidential power at the cost of traditional checks and balances. In 1941, Attorney General Robert Jackson described the opinion as "a Christmas present to the President" (Jackson, 1941, p. 201). Executive branch attorneys cite the decision with great frequency. As noted by Harold Koh, Justice Sutherland's "lavish description of the president's powers is so often quoted that it has come to be known as the '*Curtiss Wright*, so I'm right' cite—a statement of deference to the president so sweeping as to be worthy of frequent citation in any government foreign-affairs brief" (Koh, 1990, p. 94).

LITIGATION ON THE SOLE-ORGAN DOCTRINE

Because of a challenge to legislation passed by Congress in 2002, the validity of the sole-organ doctrine would be reexamined in federal court. In signing a bill that year, President George W. Bush objected to language in

Section 214(d). It stated that for purposes of registration of birth certification of nationality, or issuance of a passport of a US citizen born in Jerusalem, the Secretary of State "shall, upon the request of the citizen or the citizen's legal guardian, record the place of birth as Israel" (116 Stat. 1366). Bush (2002) stated that if this language were interpreted to impose a legislative requirement, it would "impermissibly interfere with the President's constitutional authority to formulate the position of the United States, speak for the Nation in international affairs, and determine the terms on which recognition is given to foreign states" (p. 1698). The language "speak for the Nation" was an apparent allusion to John Marshall's sole-organ speech in 1800 and its subsequent misinterpretation by the Supreme Court in *Curtiss-Wright.*

The constitutionality of Section 214(d) was challenged in court. On July 23, 2013, the D.C. Circuit relied five times on the sole-organ doctrine to hold that this statutory provision "impermissibly infringes" on the President's power to recognize foreign governments (*Zivotofsky v. Kerry*, 2013). The court acknowledged that the sole-organ doctrine was dicta in the *Curtiss-Wright* decision, but to the D.C. Circuit it was Supreme Court dicta and therefore controlling on lower courts. There was no recognition by the D.C. Circuit that the dicta had been regularly challenged by scholars as erroneous.

In response to the decision by the D.C. Circuit, I filed an amicus brief with the Supreme Court on July 17, 2014, asking it to correct the errors in *Curtiss-Wright* that had expanded presidential power in external affairs and damaged the system of checks and balances (Fisher, 2014). My brief identified the false assertions in *Curtiss-Wright*, reviewed scholarly evaluations of *Curtiss-Wright*'s dicta, and urged the court to "take steps to correct the erroneous dicta that appear in *Curtiss-Wright*, errors that have misguided federal courts, the Justice Department, Congress, some scholarly studies, and the general public" (p. 35). When the court is in session, the *National Law Journal* each week selects a brief that merits attention. On November 3, 2014, it selected mine, featuring this heading: "Can the Supreme Court Correct Erroneous Dicta?" (Schuman, 2014).

In September 2014, a brief filed by the Justice Department with the Supreme Court in *Zivotofsky* stated: "The principle that the Nation must speak with one voice in foreign affairs, see *United States v. Curtiss-Wright Exp. Corp.*, 299 U.S. 304, 319–320 (1936), therefore applies with particular force to recognition decisions" (U.S. Justice Department, 2014, p. 9). In a subsection called "The Reception Clause confers recognition powers on the President," the department noted: "The primary source of the President's recognition power is Article II's grant of authority to the President alone to 'receive Ambassadors and other public Ministers,' U.S. Const. Art. II, § 3" (p. 13). The recognition power is thus implied, not expressly stated. Congress has an implied power to make passport policy. The task before the Supreme Court was to resolve two competing implied powers, not announce plenary and exclusive powers for the President.

Jettisoning the Sole-Organ Doctrine

Writing for the court, Justice Anthony Kennedy reviewed the position offered by Secretary of State John Kerry, who urged the court to define executive power over foreign affairs in broad terms, relying on language in *Curtiss-Wright* describing the President as "the sole organ of the federal government in the field of international relations" (*Zivotofsky v. Kerry*, 2015, p. 2089). In response, the court said it "declines to acknowledge that unbounded power. … The *Curtiss-Wright* case does not extend so far as the Secretary suggests" (p. 2089). After 79 years, the court had finally discarded the erroneous dicta about the President as "sole organ" in external affairs.

As one reads further into the decision, the court did not make a clean break with *Curtiss-Wright*. It gave fresh life to the erroneous dicta in *Curtiss-Wright* by claiming that the President "has the sole power to negotiate treaties" (p. 2086). It did not acknowledge that when the D.C. Circuit in *Zivotofsky* upheld presidential power it relied five times on the sole-organ doctrine. Nor did the court explain how Justice Sutherland misrepresented John Marshall's speech in 1800. Did the court consider it inappropriate to point an accusing finger at a particular Justice and underscore the failure of his colleagues to double-check Marshall's language to make sure it was being properly cited? Would that explanation discredit the Supreme Court as an institution capable of reliable constitutional analysis? At no time did the court cite scholarly articles that from 1938 to the present time identified the various errors in *Curtiss-Wright* about the existence of plenary and exclusive powers for the President in external affairs.

More seriously, the court created a new model to support independent presidential power in external affairs. It relied in part on language that Alexander Hamilton included in one of the Federalist Papers written to support ratification of the US Constitution. In Federalist No. 70, he emphasized that with "unity" in the President comes four other qualities enabling the President to exercise "to a greater degree, '[d]ecision, activity, secrecy, and dispatch'" (p. 2086). The court identified those four qualities as though they are always salutary, meriting public, congressional, and judicial support for presidential initiatives in foreign affairs. It failed to understand that Presidents who act pursuant to those qualities have done great damage to constitutional government.

First, the court did not acknowledge that the qualities of unity, decision, activity, secrecy, and dispatch accurately describe the governments of monarchs and dictators. Second, the court did a pick and choose, highlighting Federalist No. 70 while ignoring Hamilton's Federalist No. 75 that warned against concentrating foreign affairs in the President. In analyzing the treaty power, he said that several writers placed that power "in the class of executive authorities, yet this is evidently an arbitrary disposition; for if we attend carefully to its operation, it will be found to partake more of the legislative than

of the executive character" (Wright, 2002, p. 476). Hamilton then issued a very stern warning: "The history of human conduct does not warrant that exalted opinion of human virtue which would make it wise in a nation to commit interests of so delicate and momentous a kind, as those which concern its intercourse with the rest of the world, to the sole disposal of a magistrate created and circumstanced as would be a President of the United States" (p. 477). Ignoring that language underscores the Supreme Court's bias in favor of independent presidential power in external affairs.

Third, the court took no note of how presidential actions abroad have done substantial damage to the country, its Constitution, and their own record in office. Consider President Truman's decision to allow US troops in South Korea to travel northward toward Manchuria, prompting Chinese troops to enter in great numbers. His initiative fully incorporated the qualities of unity, decision, activity, secrecy, and dispatch, but the result was a costly stalemate with heavy losses among US and allied forces, discrediting Truman's record in office (Fisher, 2017, pp. 112–118). Recall the decision of President Johnson to escalate the war in Vietnam, again with great costs to the US and to Johnson personally (pp. 152–158). Throughout the escalation, Johnson relied on the qualities of unity, decision, activity, secrecy, and dispatch. Other examples are as follows: President Reagan's involvement in Iran-Contra, President George W. Bush using military force against Iraq on the basis of six claims that Saddam Hussein possessed WMDs (six claims found to be entirely empty), and President Obama ordering military action against Libya in 2011, leaving behind a country broken legally, economically, and politically (pp. 208–213, 262–273, 287–291).

Three Justices in *Zivotofsky* identified serious deficiencies in the court's ruling. Justice Scalia, in a dissent joined by Chief Justice Roberts and Justice Alito, objected that the decision "does not rest on text or history or precedent." Instead, the majority relied on "functional considerations," such as the assertion that the nation "must speak with one voice" about the status of Jerusalem (*Zivotofsky*, p. 2123). To Scalia, the "vices of this mode of analysis go beyond mere lack of footing in the Constitution. Functionalism of the sort the court practices today will *systematically* favor the unitary President over the plural Congress in disputes involving foreign affairs" (p. 2123, emphasis in original).

In a dissent joined by Justice Alito, Chief Justice Roberts charged that the court's decision "is a first: Never before has this court accepted a President's direct defiance of an Act of Congress in the field of foreign affairs" (p. 2113). For the first 225 years, "no President prevailed when contradicting a statute in the field of foreign affairs" (p. 2113). Roberts found the court's textual basis as "even more tenuous. The President does have power to make treaties and appoint ambassadors. Art. II, § 2. But those authorities are *shared* with Congress, ibid., so they hardly support an inference that the recognition

power is *exclusive*" (p. 2114, emphasis in original). Roberts correctly pointed out that *Curtiss-Wright* "did not involve a claim that the Executive could contravene a statute; it held only that he could act pursuant to a legislative delegation" (p. 2115).

The court in *Zivotofsky* appeared to accept the general pattern of some scholars and the media to lionize the American President and embrace heroic properties, including the capacity to act instinctively and reliably for the national interest (Fisher, 2012). In upholding for the first time an exclusive authority of the President to recognize foreign governments, the court in *Zivotofsky* relied on broad, abstract, and careless language in describing executive power. An article in the *Harvard Law Review* by Jack Goldsmith (2015) said that until the Supreme Court's decision in *Zivotofsky*, executive branch lawyers "had to rely on shards of judicial dicta," but now had a Supreme Court ruling "with broad arguments for presidential exclusivity in a case that holds that the President can ignore a foreign relations statute" (p. 114). To Goldsmith, the court "simply bootstrapped poor textual and structural arguments for an uncontested independent presidential power into conclusive arguments for a different-in-kind exclusive presidential power—without even looking at Article I" (p. 122). There should be little doubt, he said, that executive branch lawyers will exploit the court's "untidy reasoning" and interpret its "pro-executive elements for all they're worth" (p. 146).

It is quite remarkable that the Supreme Court in 1936 would commit plain error, even if in dicta, and allow the error to expand presidential power decade after decade despite scholars regularly identifying the error and explaining why John Marshall never promoted "plenary and exclusive" power for the President in external affairs. Scholars not only identified the court's misinterpretation of his sole-organ speech but also other errors claiming that the President possessed exclusive authority to negotiate treaties and that somehow sovereignty passed from England directly to the American national government when in fact many of the states exercised sovereign power by entering into treaties. In jettisoning the sole-organ doctrine, the court continued to endorse the erroneous dicta about the President's treaty-negotiation power.

In attributing to the President the qualities of unity, decision, activity, secrecy, and dispatch, the court assumed with this dicta that those qualities are inherently benign, constructive, and consistent with the Constitution and self-government. It gave no thought that those same qualities have been exercised by monarchs and dictators, nor did it display any understanding that US Presidents have relied on those five qualities to do great harm to the country and its constitutional system. Throughout its decision, the court demonstrated a clear bias in favor of independent presidential power in external affairs, with little appreciation to the constitutional power accorded to Congress and the crucial importance of checks and balances.

References

Adler, D. G. (1988). The constitution and presidential warmaking: The enduring debate. *Political Science Quarterly, 103,* 1–36.

Bas v. Tingy, 4 U.S. 37 (1800).

Blackstone, W. (1765). *Commentaries on the laws of England*. Oxford: Clarendon Press.

Bruff, H. H. (2015). *Untrodden ground: How presidents interpret the Constitution*. Chicago: University of Chicago Press.

Bush, G. W. (2002, September 30). *Public papers of the presidents of the United States,* p. 1697.

Cardozo, B. N. (1921). *The nature of the judicial process*. New Haven and London: Yale University Press.

Durand v. Hollins, 8 Fed. Cas. 111 (Cir. Ct. S.D.N.Y. 1860) (Case No. 4,186).

Farrand, M. (1966). *The records of the Federal Convention of 1787*. New Haven and London: Yale University Press.

Fisher, L. (1989). Congressional participation in the treaty process. *University of Pennsylvania Law Review, 137,* 1511–1522.

Fisher, L. (2012). Teaching the presidency. *PS: Political Science & Politics, 45*, 17–31. Available at http://www.loufisher.org/docs/ci/teach.pdf.

Fisher, L. (2014). Brief amicus curiae of Louis Fisher in support of petitioner, U.S. Supreme Court, *Zivotofsky v, Kerry*, No. 13-628, July 17, 2014. Available at http://www.loufisher.org/docs/pip/Zivotofsky.pdf.

Fisher, L. (2016). The staying power of erroneous dicta: From *Curtiss-Wright* to *Zivotofsky. Constitutional Commentary, 31,* 149–219.

Fisher, L. (2017). *Supreme Court expansion of presidential power: Unconstitutional leanings*. Lawrence: University Press of Kansas.

Glennon, M. J. (1988). Two views of presidential foreign affairs power: *Little v. Barreme* or *Curtiss-Wright? Yale Journal of International Law, 13,* 5–20.

Goebel, J., Jr. (1938). Constitutional history and constitutional law. *Columbia Law Review, 38,* 555–577.

Goldsmith, J. (2015). *Zivotofsky II* as precedent in the executive branch. *Harvard Law Review, 129,* 112–146.

Jackson, R. H. (1941). *The struggle for judicial supremacy: A study of a crisis in American power politics*. New York: Vintage Books.

Jackson, R. H. (1945). Full faith and credit—The lawyer's clause of the Constitution. *Columbia Law Review, 45,* 1–34.

Kelly, A. H. (1965). Clio and the court: An illicit love affair. *Supreme Court Review, 1965*, 119–158.

Keynes, E. (1982). *Undeclared war: Twilight zone of constitutional power*. University Park and London: The Pennsylvania State University Press.

Koh, H. H. (1990). *The national security Constitution: Sharing power after the Iran-Contra affair*. New Haven and London: Yale University Press.

Levitan, D. M. (1946). The foreign relations power: An analysis of Mr. Justice Sutherland's theory. *Yale Law Journal, 55,* 467–497.

Little v. Barreme, 6 U.S. (2 Cr. 170). (1804).

Locke, J. (1690). *Second treatise of civil government*. London: J. M. Dent & Sons, Ltd. and New York: E. P. Dutton & Co., Inc., 1962.

Richardson, J. D. (1897–1925). *A compilation of the messages and papers of the Presidents* (20 vols.). New York: Bureau of National Literature.

Schuman, J. (2014, November 3). Brief of the week: Can the Supreme Court correct erroneous dicta? *National Law Journal*. Available at http://www.loufisher.org/docs/pip/fisherbrief.pdf.

Shane, P. M. (2009). *Madison's nightmare: How executive power threatens American democracy*. Chicago: University of Chicago Press.

Simones, A. (1996). The reality of *Curtiss-Wright*. *Northern Illinois University Law Review, 16,* 411–431.

Sutherland, G. (1919). *Constitutional power and world affairs*. New York: Columbia University Press.

Talbot v. Seeman, 4 U.S. 1. (1801).

United States v. Smith, 27 Fed. Cas. 1192 (C.C.N.Y. 1806) (No. 16,342).

United States v. Curtis Wright Corp., 299 U.S. 304. (1936).

U.S. Justice Department. (2014). "Brief of the Respondent," *Zivotofsky v. Kerry*, U.S. Supreme Court, No. 13-628.

Wright, B. F. (2002). *The Federalist: The famous papers on the principles of American government*. New York: Metro Books.

Zivotofsky v. Kerry, 725 F.3d 197. (D.C. Cir. 2013).

Zivotofsky ex rel. Zivotofsky v. Kerry, 135 S.Ct. 2076. (2015).

Reality Monitoring in Politics

Gary D. Bond, Sara M. Schewe, Angelica Snyder
and Lassiter F. Speller

We recently completed an applied linguistic investigation into statements made in the contentious 2016 US presidential primary debates, in which we automatically coded Reality Monitoring (RM) characteristics of six candidates' statements with software to investigate the relative deceptiveness of their debate language (Bond et al., 2017). Reality Monitoring is explained in some detail in the *Reality Monitoring* section later in this chapter. We found that debate language in the primaries was as deceptive as fact-checked lies of the same candidates. In this chapter, we will assess the 2016 campaign season in finer detail to (a) try to understand in qualitative ways how and possibly why candidates were deceptive in both primary and general election debates; (b) consider how imagination can be characterized as a sibling of deception, especially in the political arena; and (c) evaluate how well the RM framework captures the relative truthfulness of fact-checked statements and debate statements from finalist candidates Hillary Clinton and Donald Trump. We look at deception in the debates and gauge imagination/deception and RM a little later, but first we talk about the major problem of finding ground truth in the exceptionally gray-area world (Bond & Speller, 2010) of political language.

G. D. Bond (✉) · S. M. Schewe · A. Snyder
Eastern New Mexico University, Portales, NM, USA
e-mail: Gary.Bond@enmu.edu

L. F. Speller
The Ohio State University, Columbus, OH, USA

953

The Quest for Ground Truth: Difficult to Find in Politics

Werner (2016) has painted an interesting picture of how fact-checkers have become gatekeepers that exercise control over the accuracy of information, at least in the Western world, since the early 2000s. We looked to fact-checkers to try to dig up ground truth on candidates' statements in our earlier work. We used politicians' fact-checked statements that had been verified by PolitiFact, one of three major US fact-checking organizations. One important problem in field studies in the deception detection domain is establishing ground truth. This problem explains in part why there are few field studies, especially in the realm of politics. Werner indicates that fact-checkers are committed to strict guidelines in reporting in an almost scientific, non-partisan way (Werner, 2016). We reviewed PolitiFact and the method behind its fact-checking process to feel reasonably safe in using that website's ground truth statements from candidates.

PolitiFact (http://www.politifact.com) fact-checks candidates' statements and categorizes them on a "Truth-O-Meter" as true, mostly true, half true, false, and "pants on fire" lies. For an extensive description of the website's procedures and principles for fact-checking and assigning Truth-O-Meter ratings, see Adair and Holan (2013). Concisely, writers research public statements rooted in verifiable facts, and a panel of at least three editors select the Truth-O-Meter ruling, according to PolitiFact's description of fact-checking procedures. Braun, Van Swol, and Vang (2015) assessed statements on the PolitiFact website in a political deception study and could reliably differentiate lies and truths based on PolitiFact's veracity categories. Given that fact-checkers are relatively scientific in their method, and fact-checked statements have an editorial peer-review system in place, and the veracity of PolitiFact's fact-checked statements have been scientifically parsed, we believe that PolitiFact is a valid source for ground truth political statements. Given that we now have a source from which a pool of ground truth statements can be had, we can use those data as a baseline against which we can compare fact-checked lies and debate statements.

Exercises in Political Truth-Stretching

Arendt (1968, 1969) has described some deceptive tactics that politicians use including those related to the construction and maintenance of image; imagining future situations, policies, and events; manipulating, denying, or hiding facts; and self-deception. Arendt's ideas allow us to describe how, and some of the reasons why politicians stretch the truth. In this section, we also take a qualitative look at statements in the primary and general election debates to underscore the ways in which politicians engage in deception.

Self-Constructions and Image

Telling the truth is not regarded as a political virtue, since politics is action-based, according to Arendt. Politicians persuasively frame evidence about themselves to fit expectations of others. Theoretical and empirical contributions have been made over the years to the deception literature that have examined self-presentational goals of deception, and those apply quite readily to the political domain. The concept of self-monitoring was coined in work by Riggio and Friedman (1982). Self-monitoring is a grouping of skills that relate to interpreting social demands; monitoring and controlling inappropriate feelings or information and expressing one's self in a positive and socially approved style. DePaulo et al. (2003) investigated self-presentation as a way to control impressions that others form about oneself. Self-presentation can be truthful or deceptive, where truthful self-presenters edit presentations but do not go beyond the bounds of truth, but liars claim identities and manage impressions and design these efforts to mislead others (DePaulo et al., 2003). Other research has found that people lie more when they have a goal to ingratiate themselves or appear competent, and men lie more to appear competent by lying about achievements and plans (Feldman, Forrest, & Happ, 2002). Feldman et al. (2002) posited a mediated relationship in which "self-presentational style increases the demand to engage in impression management, and, in turn, increases the likelihood of lying" (p. 169).

In the 2016 debates, Mr. Trump made self-oriented statements and statements about achievements and plans. For example, in a primary debate, he said,

> I got a call from my sister and brother tonight and they said we had no idea
> Dad gave you 200 million dollars—believe me I started off with one million
> dollars—I built a company that's worth more than 10 billion dollars—and I say
> it not in a bragging way but that's the kind of thinking we need.

Kessler (2016), on the *Washington Post* fact-checking website, indicated that although Mr. Trump did benefit from a one million dollar trust his father set up for him, he also borrowed ("in the 9 million dollar range") from his father's estate. Mr. Trump seemed to be managing his image by stating that he started his business empire with a smaller amount than others claimed, a sort of "rags-to-riches" self-presentation (although even one million dollars may not be equivalent to "rags" to many).

Self-presentation, for candidate Hillary Clinton in a general election debate, included an emphasis on her record of public service and her background in law:

> When I was first lady I worked with Democrats and Republicans to create the
> Children's Health Insurance Program…I have tried my entire life to do what
> I can to support children and families…right out of law school I went to work

for the children's defense fund. Donald talks a lot about the 30 years I've been in public service. I'm proud of that. You know, I started off as a young lawyer working against discrimination against African-American children and schools and in the criminal justice system.

Ms. Clinton was fact-checked by PolitiFact on the first phrase, being "key" in helping to create the Children's Health Insurance Program, and the check indicated that it was largely accurate. Looking at the overall context of the second general election debate, which was held two days after there was news of a leaked tape of Mr. Trump's "lewd remarks in 2005 about hitting on women" (Qiu, 2016), it seems that Ms. Clinton's self-presentational strategy was to frame herself as a longtime champion of women, children, families, and the oppressed in American society. This could be considered self-presentational as she may have edited her presentations but perhaps did not go beyond the bounds of truth (DePaulo et al., 2003). She apparently tried to present a stark contrast of her image (i.e., 30 years of public service dedicated to the protection of women, children, and the oppressed) to that of Mr. Trump's implied image (i.e., implied "womanizer").

Imagination

Secondly, politicians imagine a different world and wish to place their proposed actions in place of previous actions that need to be removed. For example, Ms. Clinton said in the first general election debate that:

> what I believe is the more we can do for the middle class, the more we can invest in you, your education, your skills, your future, the better we will be off and the better we'll grow—that's the kind of economy I want us to see again.

While this might have been the imagined ideal to Ms. Clinton and her supporters, these ideas are not current policies (i.e., more educational opportunities and skills training programs for those in the middle class). Ms. Clinton's tactic was to persuade voters by offering a rosy future in which the middle classes would benefit by her policy focus on their training and skills.

In the first general election debate, Mr. Trump also imagined a different world in which cutting taxes would lead to more job creation:

> Ford is leaving—you see that their small car division leaving, thousands of jobs leaving Michigan, leaving Ohio—they're all leaving and we can't allow it to happen anymore—under my plan I'll be reducing taxes tremendously from 35 percent to 15 percent for companies, small and big businesses—that's going to be a job creator like we haven't seen since Ronald Reagan—it's going to be a beautiful thing to watch.

Imagining a future in which Americans would pay lower taxes and have greater chances to be employed was Mr. Trump's tactic to win more voters to his "side."

Arendt ([1969](#)) indicated that the construction of imagined futures is a form of political deception. We argue that imagination is a close sibling of deception, even a conjoined twin, with no mutually exclusive border. Candidates' statements include imagined outcomes that would occur in the future to win and maintain support. Imagination and deception both deviate from truth by suggesting a false, alternative, or not-experienced reality. One must have access to the contents of one's own memory as a baseline representing that which is "true," and any omission, alteration, selective sampling, or recombination of those contents of memory would be a deviation of that which is "true" to the individual.

In a review of literature in 1910, Perky indicated that imagination was usually not a separate topic in psychology texts, but was "almost always [mentioned] in close proximity to Memory" (p. 425). Psychologists then were prescient: episodic memory is sampled when one imagines past or future events (Szpunar, Watson, & McDermott, 2007). Schacter and Addis (2009) report the recruitment of overlapping medial temporal lobe functions in both activities. A reactivation of neural regions that originally processed sensory-perceptual and contextual details occurs when one remembers past true events (Danker & Anderson, 2010). Imagined future events are placed in familiar visuospatial contexts (Szpunar et al., 2007). Addis, Pan, Vu, Laiser, and Schacter (2008) found a common atemporal imagining subsystem (i.e., imagining the future and the past: anterior hippocampus, extensive medial prefrontal, and parietal regions). Important differences in remembering real events and imagining are parsed by the relative strength of contextual information (Johnson, Foley, Suengas, & Raye, 1988). Thus, imagination and lying both involve the recruitment of a patchwork of experienced events to create events for any time period (past, present, future).

Statements based in imagination can be similar to lying. Deception has been conceptualized as false communication that benefits the communicator (Mitchell, 1986); and as a deliberate attempt, which may be successful or not "to create in another a belief that the communicator considers to be untrue" (Vrij, 2000, p. 6). Imaginative communications are false, and could benefit the communicator, as in candidates' debate statements that construct imagined pasts (e.g., an imagined reconstruction that creates a brighter time than may have been, such as in Mr. Trump's characterization of Ronald Reagan's job creation record), presents (e.g., Ms. Clinton's implication that the middle class had few current benefits for education and training), and futures (e.g., Mr. Trump's imagined world in which taxes would be much lower and jobs would be plentiful). Candidates can deliberately omit, alter, sample, or recombine facts or events to bolster success in their candidacies.

Manipulating, Denying, or Hiding Facts

Arendt (1969) indicated that facts are different from opinions, and so should not be able to be disputed. One removes facts from political debate by lying about them. For example, when candidate Ted Cruz was asked in a

Republican primary debate if President Obama should nominate a successor for the late Supreme Court judge Scalia, Mr. Cruz stated, "we have 80 years of precedent of not confirming Supreme Court justices in an election year." The moderator of that debate, John Dickerson, immediately fact-checked Mr. Cruz and responded that Justice Anthony "Kennedy was confirmed in '88" (Wade, 2016). Another example of lying about facts comes from Ms. Clinton in the Detroit, Michigan general election debate: "back in the Great Recession…Donald Trump said rescuing the auto industry didn't really matter very much. He said, and I quote again, 'Let it go'"; however, FactCheck.org indicated that Mr. Trump had made several statements at the time in favor of rescuing the auto industry from bankruptcy (Robertson, 2016).

Candidate Bernie Sanders selectively omitted and twisted facts when talking about Ms. Clinton's primary wins up to the 9th Democratic debate:

> Secretary Clinton cleaned our clock in the Deep South. No question about it. We got murdered there. That is the most conservative part of this great country. That's the fact. But you know what? We're out of the Deep South now.

FactCheck.org indicated that Ms. Clinton had not only won state primaries in the southern US, but also mid-western states, western states, and Massachusetts up to the time of Mr. Sanders's statement. Further, the fact-check indicated that large African–American voter turnout, and not conservative turnout, helped her win in the southern United States (Kiely, Jackson, Farley, Robertson, & Gore, 2016). Thus, facts can be discredited by politicians by selectively omitting parts of facts, and lying about their origin, their circumstances, or their existence.

Self-Deception

Arendt (1969) suggested that politicians become successful in lying to greater numbers of people, and become invested in believing their own lies. A politician's self-deception is the "best way to deceive others…a charismatic leader is persuasive in proportion to his convictions and faith; a cynical, self-interested liar is more easily detected and can hardly become a charismatic leader" (Galeotti, 2015, p. 891). Further, it is possible that self-deception simplifies deceiving others (Trivers, 2011), because self-deceivers are less likely to exhibit deceptive cues to others (von Hippel & Trivers, 2011) and because highly confident and over-confident people who are self-deceivers are over-rated by others (Lambda & Nityananda, 2014). In a 2013 paper, Triandis asserted that if a high percentage of information that people use to construct the world comes from inside the body (which would include hopes, imagined attractive objects, prejudices, stereotypes, theories, etc.), then there may be a higher probability they will engage in self-deception. This may partially explain embedding of biased or false beliefs about others and the world.

Biased belief formation can result from a motivation to gather evidence (Mele, 2008). A self-deceiver selectively contemplates data that are confirmatory and neglects data to the contrary; confirmatory data are more vivid to the self-deceiver than those data that are disconfirming. These may be reasons why politicians fabricate information about facts, policies, and themselves, out of "hoped-for" fantasies rather than reality.

REALITY MONITORING

We will now describe a memory framework that has experienced a metamorphosis of sorts from its original form in the 1980s, in that it has been modified and applied to the deception detection field to assess veracity of verbal content of statements: Reality Monitoring (RM). How well can the modified RM framework capture the relative truthfulness of politicians' fact-checked statements and debate statements? Originally, RM was introduced by Johnson and Raye (1981) to explain qualitatively how one can differentiate between real and imagined memories (i.e., externally generated; internally generated, respectively). Internally generated memories show greater evidence of cognitive processes and externally based memories have greater sensory-perceptual and contextual information (temporal, spatial, affective contexts). In work conducted a few years after Johnson and Raye introduced their theory, differences using some of the RM features were found between real memories and memories that were generated from false or misleading information (Schooler, Clark, & Loftus, 1988; Schooler, Gerhard, & Loftus, 1986). Subsequently, the framework was used in a deception paradigm to successfully discriminate between other people's truths and lies (Alonso-Quecuty, 1992). Later, Sporer (1997) applied RM criteria that could be rated by expert judges to discern between truthful and deceptive statements. He used eight subscales of the Judgment of Memory Characteristics Questionnaire (JMCQ; Sporer & Kuepper, 1995), which were derived from the Memory Characteristics Questionnaire (MCQ; Johnson et al., 1988). The MCQ was created to test an individual's RM ratings of his or her own memories; and the JMCQ was modified to allow for an observer to judge other people's accounts based on these RM criteria: clarity, sensory experiences, spatial and time information, emotions and feelings, reconstructability of the story, realism, and cognitive operations (Sporer, 1997). Sporer found 75% classification for self-experienced accounts and 67.5% classification of invented accounts using the JMCQ criteria.

Reviews of RM research have shown about a 70% classification rate for truth and lies over studies, but individual features of RM have varied widely in studies from Sporer's (1997) work. Further, diagnosticity of veracity from individual features varies between studies as well. In a review of 29 studies, Vrij (2008) found that people who told the truth used more sensory and contextual information, but affect and cognitive operations were

not diagnostic of veracity. Masip, Sporer, Garrido, and Herrero (2005) reported in a meta-analysis that perceptual processes, contextual information (including time features) and realism or plausibility of accounts distinguished statement veracity, but the authors lamented differences in procedures and operationalizations used in individual studies, which makes a uniform review of studies problematic. A more recent meta-analysis by Hauch, Blandón-Gitlin, Masip, and Sporer (2015) suffered from the same review difficulties due to differing procedures, samples, and operationalizations of RM features across studies. The authors included two RM studies and 12 "combination" studies, and found a small effect for sensory-perceptual processes, and a positive effect for cognitive processes. Overall accuracy rates based on RM classification of veracity has shown 72% for truth and 66% for lie statements over 10 studies and when summing scores of different criteria, average RM accuracy is 69% (Vrij, 2008).

Linguistic Inquiry and Word Count (LIWC) RM Coding

Trained coders are primarily used in RM deception detection studies, where visual and auditory details, temporal and spatial details, and cognitive operations are counted (e.g., Gnisci, Caso, & Vrij, 2010). A small number of RM studies by Bond and colleagues (Bond & Lee, 2005; Bond & Speller, 2010; Bond et al., 2017), and by Vrij, Mann, Kristen, and Fisher (2007), used Linguistic Inquiry and Word Count (LIWC; Pennebaker, Booth, Boyd, & Francis, 2015) as a relatively automatic alternative to human coders. In addition to many other psychological categories, LIWC generates percentages of dictionary-coded categories for sensory-perceptual, temporal, spatial, affective, and cognitive processes. In the Bond research, perception words + space words + time words + affect words − cognitive process words were used in LIWC to produce an overall RM score, but in the Vrij et al. (2007) work, affect words were not included.

There has been skepticism by scholars comparing the effectiveness of using LIWC to code RM categories vs. human RM coders. In a meta-analysis of a small number of LIWC RM studies conducted by Hauch et al. (2015) a small effect for sensory-perceptual processes and a positive effect for cognitive processes (counter to RM memory theory) was found. Vrij et al. (2007) questioned the appropriateness of using LIWC to code RM: "although the LIWC categories may resemble the RM categories, they are not developed on the basis of RM theory" (p. 502). Vrij et al. (2007) noted differences between LIWC and human coders' assessment of the cognitive mechanisms category, where "I think she had dark hair" would be assigned as a cognitive process in the LIWC 2001 and 2007 dictionaries, but that phrase would not be coded as a cognitive process by human coders. Bond and Lee (2005), using prisoners' statements, found good classification rates for veracity in the RM framework using LIWC but spatial words were higher in lies than truth. We are interested in determining how well the RM algorithm in LIWC can capture veracity of fact-checked statements and debate statements from 2016 finalist candidates Hillary Clinton and Donald Trump in the political realm. We

will briefly summarize results of our previous investigation of RM/LIWC in the primary season (Bond et al., 2017), and then assess statements by Ms. Clinton and Mr. Trump in their faceoff in the general election.

Primary Debates

Method

Statements included in the primary election study were from Donald Trump, Hillary Clinton, Marco Rubio, Bernie Sanders, Ted Cruz, and John Kasich, who were the presidential candidates who participated in all primary debates. Candidate ages ranged from 45 to 74. The PolitFact website (http://www.politifact.com/) was used to save fact-checked truth and lie statements for those candidates. One hundred forty-eight truth and 100 lie statements were included ($n = 248$ total). Average word count for truth statements was 18.2 (SD = 9.1), and average word count for lies was 18.2 (SD = 11.7) after random deletion of words in deceptive statements to match word counts in truth statements (refer to Method in Bond et al., 2017 for rationale and procedure). Transcripts from Presidential candidates' statements in debates were found on the website http://www.presidency.ucsb.edu/debates.php, and each transcript was saved as a separate text file ($n = 63$ files).

Results

A t-test supported the effectiveness of our algorithm in LIWC (percepts + time + space + affect - cogproc words = RM score). RM scores were significantly higher in fact-checked truth ($M = 11.21$, SD = 11.3) when compared to lie statements ($M = 6.87$, SD = 15.7), $t(246) = 2.53$, $p = 0.01$, $d = 0.32$. Higher RM scores suggest greater truthfulness and lower scores imply greater use of deception. Classification rates for truths and lies were obtained in a binary logistic regression with individual RM features predicting veracity (truth, lie). The model with all predictors showed $x^2(N = 5) = 24.80$, $p = 0.0001$, $-2 \log \text{likelihood} = 309.65$, Cox & Snell $R^2 = 0.10$. Significant predictors of veracity were cognitive processes ($\beta = -0.06$, $p = 0.0007$) and perceptual processes ($\beta = -0.14$, $p = 0.0012$; in the wrong predicted direction). Affect approached significance in the model ($\beta = 0.55$, $p = 0.06$). Space and time features were not significant predictors in the model. Classification rates were 85.14% for truth statements and 40% for lie statements.

A repeated measures ANOVA was conducted with Candidate (Clinton, Cruz, Kasich, Rubio, Sanders, and Trump) as the independent variable measured on Veracity (RM scores in Lie, Truth, and Debate statements). A significant effect for Veracity, $F(1.642, 252.91) = 7.83$, $p = 0.001$, $\eta_p^2 = 0.048$, and an interaction between Veracity and Candidate, $F(8.211, 252.91) = 3.29$, $p = 0.001$, $\eta_p^2 = 0.096$ were found. Comparisons for Veracity revealed differences between Truth

$(M = 11.21, \text{SD} = 11.3)$ and Lie $(M = 6.87, \text{SD} = 15.7, d = 0.32)$ RM scores and Truth and Debate $(M = 6.41, \text{SD} = 1.3, d = 0.59)$ RM scores. RM scores for debate statements were as deceptive as lies. The outcome suggests that when researchers use the RM algorithm to detect the veracity of language, politicians have a high RM score when speaking the truth, and low RM scores when telling lies or making statements in debates.

GENERAL ELECTION DEBATES

Method

Finalist candidates for the general election were Hillary Clinton, Democrat (age 68), and Donald Trump, Republican (age 70). Their general election debate statements were found on the website http://www.presidency.ucsb.edu/debates.php. Statements from Ms. Clinton and Mr. Trump were saved as text files from each of the three general election debates ($n = 212$ total statements for Ms. Clinton; $n = 303$ total for Mr. Trump; $n = 515$ total). Total word count for Ms. Clinton was 19,436 (M statement length 91.68 words, SD = 116.1), and for Mr. Trump, total word count was 27,818 (M statement length 71.33 words, SD = 109).

Forty-nine statements were randomly selected for Ms. Clinton (31 truths, 18 lies) and 39 were saved for Mr. Trump (20 truths, 19 lies) from the PolitiFact website. Average word count for Ms. Clinton's truth statements was 18.53 (SD = 7.7), and for Mr. Trump, $M_{word\ count} = 11.42$ (SD = 6.0). For lie statements, $M_{word\ count} = 17.68$ (SD = 9.0) for Ms. Clinton, and $M_{word\ count} = 20.74$ (SD = 9.4) for Mr. Trump.

Results

Table 49.1 shows means for overall RM scores as well as individual features for each finalist candidate across lie, truth, and debate statements.

Table 49.1 Mean RM scores and features for Ms. Clinton and Mr. Trump for lie, truth, and debate statements

Feature	Hillary Clinton			Donald Trump		
	Lie	Truth	Debate	Lie	Truth	Debate
Perceptual processes	2.54	1.78	1.79	3.98[c]	1.01[d]	2.55
Space	8.40	8.03	5.26	8.44	6.10	5.09
Time	3.58	4.48	3.75	4.67	5.77	3.22
Affect	0.71[a]	4.86[b]	7.21[b]	3.45[c]	6.26	9.16[d]
Cognitive processes	15.43[a]	8.67[b]	13.19[a]	6.32[c]	9.24[c]	13.78[d]
RM score	−0.21[a]	10.48[b]	4.81[a]	14.21[c]	9.90[c]	6.24[d]

Note Superscripts that are different from each other indicate a significant difference between means
Superscripts [a] and [b] are used for Ms. Clinton, and [c] and [d] are used for Mr. Trump
RM Score derived from the algorithm perceptual processes + space + time + affect − cognitive processes

Candidate (Clinton, Trump) was the independent variable, with repeated measures on Statement Type (Truth, Lie, Debate). There was a significant mean difference between Candidates on Statement Type for RM scores, $F(1, 35) = 9.78$, $p = 0.004$, $\eta^2 = 0.28$, and the Candidate X Statement Type interaction was not significant, $F(2, 1.797) = 2.27$, $p = 0.11$. Ms. Clinton's RM scores followed the predicted RM pattern where her lie RM score was significantly lower than her truth RM score (and there was no difference between her debate and lie RM scores); however, Mr. Trump's lie RM score was not different from his truth RM score.

Discussion

A real-life, high-stakes, contentious 2016 US election process seemed to provide a perfect opportunity to describe and explain important aspects of political deception. We sought to assess the campaign season to try to understand in qualitative ways how and possibly why candidates deceive; how imagination can be characterized as a sibling of deception; and how well the RM framework captures veracity of fact-checked and debate statements from primary and finalist candidates.

Politicians engage in deception. Arendt (1968, 1969) provided a framework of deceptive tactics that politicians use to deceive. Image is extremely important to politicians; they deceive to construct and maintain image in their pursuit to become charismatic leaders (Galeotti, 2015) that appeal to potential voters. Self-presentation (DePaulo et al., 2003) is relevant to image construction and maintenance: while we afforded a few examples of how Ms. Clinton and Mr. Trump self-presented to potential voters in debates, we came across many examples of *image destruction*, which is a tactic used to build up one's own image and tear down other candidates' images. We know that Mr. Trump, who became President Trump, used image destruction by marking other candidates with degrading labels: "Crooked Hillary," "'Lyin' Ted," "Little Marco," and several others. Ms. Clinton also used image destruction tactics in reactions to news that centered on Mr. Trump. For example, the "private" video that was shown by media sources that depicted Mr. Trump talking with another person about women became an opportunity for Ms. Clinton to capitalize on in debates. How image destruction can be deceptive and how it relates to self-presentation is an interesting future direction for scholars.

Another future direction is to assess how imagination is similar to and different from deception. Politicians imagine future situations, policies, and events, and Arendt suggests that these are exercises in deception. The rosiest imagined future for the greatest number of people wins votes and brings voters into the fold, although those imagined future events may never be realized. The RM framework originally proposed the qualitative parsing of an individual's true and imagined past events. Are imagining past events qualitatively different than imagining future events?

Politicians manipulate, deny, and hide facts to deceive. This set of deceptive tactics is squarely in the province of fact-checkers. We reliably distinguished true statements of fact from deceptive statements using RM based on a fact-checking website, with two exceptions (Mr. Cruz and Mr. Trump), as we will discuss in a moment. Two interesting future directions in research would be to investigate potential biases of political fact-checkers: are they biased in fact-checking certain candidate's statements only if they think they are deceptive statements, and do fact-checkers only check facts presented by favored candidates? We also scratched the surface of politicians' self-deceptions, but the difficulty in studying self-deception and providing examples, is to sample a person's thoughts and beliefs from day to day, and then try to determine what information they are biased toward and which data they discount (Mele, 2008) for any false beliefs they eventually construct about themselves.

Finally, let us attempt to evaluate the RM framework used in an algorithm in LIWC to assess statement authenticity. The framework seemed to work relatively well (~70% classification rate) to discriminate fact-checked lies and truths and to place debate statements on a veracity continuum in the primary portion of the election. There are several weaknesses in this method that we discovered, however. Classification was very good for truthful fact-checked statements (87.8%), but below the flip of a coin (44%) for fact-checked lie statements. In terms of individual features or criteria in the RM coding, we found that cognitive process words were an important predictor of deception, and affect information approached being a significant predictor of truth; higher percentages of cognitive process words represented the single RM feature that predicts deception and higher percentages of affect words exemplified one of four features that predict truth in the RM framework. Perceptual process words were a significant predictor of veracity but in the wrong theoretical direction (indicative of deception rather than truth).

Individual features found to be significant predictors in this research deviate from findings in past RM meta-analyses, possibly because the context (political, high-stakes) is different from contexts that have been explored previously, primarily in laboratory studies with student participants. In Bond and Lee's (2005) work, the authors argued that prisoners, who used more spatial information in lie statements, were highly cognizant of being locked away in their restricted, confined spaces, and in the act of lying, possibly created deceptive statements that sought to transcend their restricted spatial context. It is likely that LIWC features vary with physical and cognitive contexts. For example, if we were to have people talk about the loss of loved ones in a research study (e.g., Bond et al., 2015), we might find artifacts in LIWC with high numbers of spatial words ("up" in Heaven; watching "over" me; "in" a better place) that reflect that cognitive context.

Importantly, cognitive process words strongly predicted deception in this research. Advance preparation for planned political deception may explain greater evidence of cognitive processes in fact-checked lie statements; and advance preparation for debates, combined with the control of

thought processes while appraising and scrutinizing responses of others in the debate (Spence, 2004) taxes cognitive processes. Affect words were part of the context of truth statements and were lower in lie and debate statements. Perceptual process words were higher in deceptive statements and in debate statements, which could possibly reflect imaginative processing, which as we asserted involves the recruitment of a patchwork of experienced events (with sensory-perceptual information) to create fiction. When politicians tell the truth, they often talk about facts, policies, plans, third-party information, and other topics that would probably not include much sensory-perceptual information. This again speaks to cognitive context—linguistic content is restricted to certain spheres of topics leaving less linguistic space to talk about autobiographical memories per se, and this is a cognitive context difference between politics and other domains.

Fact-checked samples and statement word counts from the fact-checking website were both relatively small. In the Bond et al. (2017) study, Mr. Cruz's lie RM score was higher than his truth RM score, but in PolitiFact, there were only eight fact-checked truth scores available for Mr. Cruz; hardly a sample with which to be comfortable. There were no differences between Mr. Trump's RM scores for truth and fiction. Sample size was larger for his fact-checked statements than for Mr. Cruz. However, fact-checked statements overall were relatively short in word count. Given that Vrij and others have called into question using LIWC to code RM for theoretical and coding reasons, we wonder whether short fact-checked statements could be captured better by human coders who take into account semantic context of statements. However, we did find that most candidates followed the predicted direction for RM scores (higher in truth, lower in lies).

RM in the political domain can prime us with more than a few thoughts. Fact-checkers monitor reality for the public so that statements made within the realities of candidates can be held up for scrutiny. Fact-checkers are journalistic gatekeepers in a world in which many media sources seem to be biased, and so fact-checkers are important and necessary to keep the public from being mired in biased gray areas. Political truth-twisting can be monitored with methods of statement analysis like RM and we can become significantly better at detecting deception produced by politicians. This largely depends upon the unbiased, relatively scientific peer-reviewed system that fact-checkers maintain in order to establish ground truth. It also depends upon using the best method for RM coding, whether it be with human coders or using an algorithm in a computer software program. Lastly, when we monitor the reality of debate statements, we can be relatively sure that most of the language produced will be deceptive. Debates are usually prepared for and simulated in detail by participants. Debates will most certainly include exercises in truth-stretching, where images are firmed up (and others' images are dented or destroyed), imagined future policies and programs or imagined altered past events and records are effortlessly uttered, and altered or inaccurate facts are bandied about on the political stage.

References

Adair, B., & Holan, A. D. (2013, November 1). The principles of PolitiFact, PunditFact and the Truth-O-Meter. *PolitiFact*. Retrieved from http://www.politi-fact.com/truth-o-meter/article/2013/nov/01/principles-politifact-punditfact-and-truth-o-meter/.

Addis, D. R., Pan, L., Vu, M. A., Laiser, N., & Schacter, D. L. (2008). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia, 14*(11), 2222–2238. https://doi.org/10.1016/j.neuropsychologia.2008.10.026.

Alonso-Quecuty, M. (1992). Deception detection and reality monitoring: A new answer to an old question? In F. Loesel, D. Bruder, & T. Blisener (Eds.), *Psychology and law: International perspectives* (pp. 228–332). Berlin: de Gruyter.

Arendt, H. (1968). Truth in politics. In H. Arendt (Ed.), *Between past and future: Eight exercises in political thought* (pp. 223–259). New York: Penguin.

Arendt, H. (1969). Lying in politics. In H. Arendt (Ed.), *Crises of the republic* (pp. 1–48). New York: Harcourt, Brace, and Company.

Bond, G. D., Holman, R. D., Eggert, J. L., Speller, L. F., Garcia, O. N., Mejia, S. C., …, Rustige, R. (2017). 'Lyin Ted,' 'Crooked Hillary,' and 'Deceptive Donald': Language of lies in the 2016 U.S. Presidential debates. *Applied Cognitive Psychology, 31*(6), 668-677. https://doi.org/10.1002/acp.3376.

Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology, 19,* 313–329.

Bond, G. D., & Speller, L. F. (2010). Gray area messages. In M. S. McGlone & M. L. Knapp (Eds.), *The interplay of truth and deception: New agendas in communication* (pp. 35–53). New York: Routledge.

Bond, G. D., Walker, W. R., Bargo, A. J. B., Bansag, M. J., Self, E. A., Henderson, D. X., …, Alderson, C. J. (2015). Fading affect bias in the Philippines: Confirmation of the FAB in positive and negative memories but not for death memories. *Applied Cognitive Psychology*, *30*, 51–60. https://doi.org/10.1002/acp.3166.

Braun, M. T., Van Swol, L. M., & Vang, L. (2015). His lips are moving: Pinocchio effect and other lexical indicators of political deceptions. *Discourse Processes, 52*(1), 1–20. https://doi.org/10.1080/0163853X.2014.942833.

Danker, J. F., & Anderson, J. R. (2010). The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin, 136*(1), 87–102. https://doi.org/10.1037/a0017937.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118.

Feldman, R. S., Forrest, J. A., & Happ, B. R. (2002). Self-presentation and verbal deception: Do self-presenters lie more? *Basic and Applied Social Psychology, 24*(2), 163–170.

Galeotti, A. E. (2015). Liars or self-deceived? Reflections on political deception. *Political Studies, 63,* 887–902. https://doi.org/10.1111/1467-9248.12122.

Gnisci, A., Caso, L., & Vrij, A. (2010). Have you made up your story? The effect of suspicion and liars' strategies on reality monitoring. *Applied Cognitive Psychology, 24,* 762–773. https://doi.org/10.1002/acp.1584.

Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to

deception. *Personality and Social Psychology Review, 19*(4), 307–342. https://doi.org/10.1177/1088868314556539.

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117,* 371–376.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67–85.

Kessler, G. (2016, September 27). Fact-check: How much help did Trump's father give his son? *Washington Post.* Retrieved from https://www.washingtonpost.com/politics/2016/live-updates/general-election/real-time-fact-checking-and-analysis-of-the-first-presidential-debate/fact-check-how-much-help-did-trumps-father-give-his-son/?utm_term=.d989f367d5ea.

Kiely, E., Jackson, B., Farley, R., Robertson, L., & Gore, D. (2016, April 15). Fact-checking the ninth Democratic debate. *FactCheck.org.* Retrieved from http://www.factcheck.org/2016/04/factchecking-the-ninth-democratic-debate/.

Lambda, S., & Nityananda, V. (2014). Self-deceived individuals are better at deceiving others. *PLoS One, 9*(8), e104562. https://doi.org/10.1371/journal.pone.0104562.

Masip, J., Sporer, S., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law, 11,* 99–122. https://doi.org/10.1080/10683160410001726356.

Mele, A. R. (2008). Self-deception and delusions. In T. Bayne & J. Fernández (Eds.), *Delusions and self deception: Affective and motivational influences on belief-formation* (pp. 55–70). Hove: Psychology Press.

Mitchell, R. W. (1986). A framework for discussing deception. In R. W. Mitchell & N. S. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 3–40). Albany: State University of New York Press.

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic inquiry and word count: LIWC2015.* Austin, TX: Pennebaker Conglomerates. www.LIWC.net.

Perky, C. W. (1910). An experimental study of imagination. *The American Journal of Psychology, 21*(3), 422–452. http://www.jstor.org/stable/1413350.

Qiu, L. (2016, October 9). Fact-checking the second presidential debate. *PolitiFact.* Retrieved from http://www.politifact.com/truth-o-meter/article/2016/oct/09/fact-checks-second-presidential-debate/.

Riggio, R. E., & Friedman, H. S. (1982). The interrelationships of self-monitoring factors, personality traits, and nonverbal social skills. *Journal of Nonverbal Behavior, 7*(1), 33–45.

Robertson, L. (2016, October 14). Clinton's auto bailout falsehood. *FactCheck.org.* Retrieved from http://www.factcheck.org/2016/10/clintons-auto-bailout-falsehood/.

Schacter, D. L., & Addis, D. R. (2009). On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philosophical Transactions of the Royal Society (B), 364,* 1245–1253. https://doi.org/10.1098/rstb.2008.0308.

Schooler, J. W., Clark, L. A., & Loftus, E. F. (1988). Knowing when memory is real. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues.* New York: Wiley.

Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 12*(2), 171–181.

Spence, S. (2004). The deceptive brain. *Journal of the Royal Society of Medicine, 97*(1), 6–9.

Sporer, S. L. (1997). The less-traveled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology, 11*, 373–397.

Sporer, S. L., & Kuepper, B. (1995). *Realitaetsueberwachung und die Buerteilung des Wahrheitsgehaltes von Erzaehlungen: Eine experimentelle Studie* [Reality monitoring and the judgment of credibility of stories: An experimental investigation]. *Zeitschrift für Sozialpsychologie, 26*, 173–193.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences United States of America, 104*(2), 642–647. https://doi.org/10.1073/pnas.0610082104.

Triandis, H. C. (2013). Self-deception: An introduction. *Acta de Investigación Psicológica, 3*(2), 1069–1078. https://doi.org/10.1016/S2007-4719(13)70952-9.

Trivers, R. (2011). *Deceit and self-deception: Fooling yourself the better to fool others.* London: Penguin.

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34*(1), 1–16. https://doi.org/10.1017/S0140525X10001354.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice.* Chichester, UK: Wiley.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities.* Chichester, UK: Wiley.

Vrij, A., Mann, S., Kristen, S., & Fisher, R. P. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*(5), 499–518. https://doi.org/10.1007/s10979-006-9066-4.

Wade, P. (2016, February 14). The three biggest lies you were told at Saturday's Republican debate. *Esquire.* Retrieved from http://www.esquire.com/news-politics/news/a42146/gop-debate-fact-check/.

Werner, H. (2016, November). *The effect of fact-checking on trust in politicians and trust in media sources during the U.S. presidential campaign 2016.* Paper presented at the WAPOR conference for Political Trust in Contemporary Representative Democracies, Barcelona, Spain.

CHAPTER 50

Propaganda, Politics and Deception

David Miller and Piers Robinson

Popular current debates surrounding "fake news" and propaganda have elevated the perceived importance of the role of deception in politics. Much of the mainstream discussion of these matters has tended to focus on the controversial US President Donald Trump, the prevalence of "fake news" across social and alternative/independent media, and a renewed fixation on the alleged Russian threat to Western interests and its mischievous employment of propaganda as a part of hybrid warfare. Trump has frequently charged mainstream US media with disseminating "fake news" whilst there is now extensive angst over the circulation of both fake news and propaganda across social media. Russia has been frequently accused of employing both cyber warfare and state-backed media outlets such as *RT* and *Sputnik* in order to influence elections across the West, in part by spreading disinformation.

However, what we now frequently refer to as "fake news" and its deployment as part of a political strategy are nothing new to Western democracies.

D. Miller (✉)
School for Policy Studies, University of Bristol, Bristol, UK
e-mail: david.miller@bristol.ac.uk

P. Robinson
School of Social Sciences, University of Sheffield, Sheffield, UK
e-mail: piers.robinson@sheffield.ac.uk; piers.robinson@propagandastudies.ac.uk

As shall be discussed in this chapter, deception as a political strategy has a long lineage, dating back to Aristotle, and has been frequently employed across Western democracies. At the same time, although we frequently associate propaganda with external hostile states, e.g. Russia, and, in doing so, disassociate it from Western liberal democracies, we do in fact live in societies in which manipulated information is ubiquitous. This chapter maps the role of both deception and propaganda in liberal democracies. The first section charts the history of deception as a political strategy, detailing in particular the rationale and ethics behind its use. The second section introduces the concept of contemporary propaganda and explains how manipulated information, frequently involving deception, has become integral to the exercise of power within Western democracies. The final section addresses key issues now emerging with respect to propaganda and deception in the contemporary media environment, characterised by Internet-based communication and the increasing employment of artificial intelligence technology designed to "persuade" and "influence" the beliefs and conduct of democratic citizenry. The nefarious and deleterious consequences of these dynamics will be highlighted.

THE LONG HISTORY OF DECEPTION IN POLITICS: FROM PLATO'S NOBLE LIE TO NEOCONSERVATIVE IDEOLOGY AND MARX'S FALSE CONSCIOUSNESS

As John Mearsheimer noted in *Why Leaders Lie*, lying and deception receive scant attention from scholars of political science and international relations. This might be because of ideological bias whereby predominantly liberal scholars simply do not perceive that cherished democratic systems might be seriously compromised by deception and lies. Or it might be because the intellectually lazy, but extremely effective, label "conspiracy theorist" is so frequently levelled at anyone who explores covert activity amongst politically powerful actors (Dentith, 2018). It might be because academics feel ill-equipped in terms of methods and research tools to penetrate sufficiently the power networks that lie behind organised "disinformation" campaigns (Miller, Brown, Dinan, & Stavinoha, forthcoming). At an epistemological level, it is highly likely that the postmodern turn, which refutes the ability to distinguish true from false, has made the question of deception a moot point. Most simply, perhaps, the difficulty of distinguishing between intentional deception, misperception and ideological frameworks causes academics to shy away from a concerted effort to analyse and pinpoint political deception (Corner, 2007).

This lacuna is odd given the long history of deception in politics. In ancient Athens, Plato's concept of the "noble lie" (360 BC, Book 3, pp. 414–415), relaying the thinking of Socrates, referred to the importance of deceptive myths that were essential to maintaining order in society. The idea here was that to ensure harmony in the context of a social hierarchy, myths needed to be created in order to help people accept their location within

that hierarchy: God made some to rule (the golden race), others to build (iron and bronze workers) and still others to fight (soldiers). Aristotle's *On Rhetoric* (2013 [230 BC]) attempted to distinguish between sophistry and rhetoric with a noble purpose although, according to Corner (2007, p. 672), his own arts of persuasion seem at times close to the advocacy of deception. Generally, Athenian thinkers such as Aeschylus, Sophocles, Thucydides, and Plato distinguished "persuasion brought about by deceit (dolos), false logic, coercion, and other forms of chicanery from persuasion (peitho)" (Lebow, 2008, p. 28) achieved through sincere dialogue. Perhaps more than any other individual thinker, sixteenth-century Niccolò Machiavelli has come to epitomise the doctrine of deception as a necessary political tactic. *The Prince* sets a rationale for the importance of deception as a political strategy. Machiavelli advised that, because men are bad, "and will not keep faith with you, you too are not bound to observe it with them" (p. 63). Importantly, the "Prince" (i.e., the person who governs) must and can: "be a great pretender and dissembler; and men are simple, and so subject to present necessities, that he who seeks to deceive will always find someone who will allow himself to be deceived" (2003 [1532], Ch. 18). Broadly speaking, Machiavelli understood deception as one amongst many tactics, others including coercion, that are necessary in order to secure and protect the state.

Ideas that deception might be a necessary political strategy whether to protect state interests or as part of an elitist ideological mindset have persisted into the contemporary era. For example, the work of Leo Strauss represents a contemporary incarnation of Plato's "noble lie" and has been associated with neoconservative thinking (Strauss, 1975; see also Strauss, 1958). Here, Plato's noble lie can be clearly seen in the idea that democratic politics is too idealistic and that the greater good can only be achieved by deferring to wise and enlightened elites. Strauss's concern is that, at times, the truth would threaten political stability and, consequently, deception becomes essential to political order and stability. The elitist and indeed anti-democratic sensibility of this perspective can be seen in the following quote from the neoconservative writer Irving Kristol:

> There are truths appropriate for children; truths that are appropriate for students; truths that are appropriate for educated adults; and truths that are appropriate for highly educated adults, and the notion that there should be one set of truths for everyone is a fallacy of modern day democracy, it simply does not work. (cited in Oborne, 2014, p. 184)

A less ideological and more practical reification of the importance of deception can be seen in the work of John Mearsheimer (2011). Informed by the realist theoretical perspective on international politics, which emphasises the inherent dangers of an anarchic international system and the importance of states to protect their own security above all else, Mearsheimer sees deception as simply a necessary part of protecting state security. Interestingly, and

counter-intuitively, he argues that interstate deception through lying is comparatively rare and that, more frequently, leaders deceive their own publics in order to defend what they perceive to be the national interest. So, for example, leaders might engage in fearmongering when they "see a threat emerging but think that they cannot make the public see the wolf at the door without resorting to a deception campaign" (Mearsheimer, 2011, p. 45). One example of this form of deception is Franklin Delano Roosevelt's lies to the American public to try to get the US involved in World War II (Dallek, 1979). Mearsheimer also discusses how lies might be used to cover up strategic failures (p. 67). Most controversially, at least from the perspective of liberal democracies, Mearsheimer argues that leaders will engage in deception when foreign policies fall short of liberal claims to be law-abiding actors who uphold high moral standards within the international system. Harking back to Plato's "noble lie", Mearsheimer also notes how nationalist myths, designed to foster social cohesion and support for the state, frequently involve lies and half-truths (p. 75).

Of course, others are far more reluctant to accept or tolerate deception and lying as a political strategy (Bok, 1999; Cliffe, Ramsay, & Bartlett, 2000; Ramsay, 2000). From a democratic perspective, deception is clearly deeply problematic. If publics are deceived by politicians and other powerful actors, it is very difficult to conceive how meaningful democratic debate can occur. As Bakir, Herring, Miller, and Robinson (2018a) argue, free and informed consent cannot be obtained under conditions in which information has been so manipulated that people do not possess sufficient information with which to rationally evaluate an issue. Deception, quite clearly, violates basic requirements of a democratic public sphere (Habermas, 1984). For those who emphasise the importance of democracy, although few of whom would rule out lying under all circumstances, it is essential that clear limits should be placed with respect to deception as a political strategy. Ramsay (2000) emphasises the corrosive impact of deception upon rational and effective policy-making:

> Because information is only available to a small number of people, this limits debate and hinders communication between those who need to know the facts in order to ensure that sound decisions are made. It also narrows the range of perspectives and opinions brought to bear on solving problems, restricts consideration of all the implications of a course of action and prevents criticism and dissenting views from being heard. (p. 37)

Similarly, Bok (1999) highlights how deception corrupts political processes creating "dangers of bias, self-harm, proliferation, and severe injuries to trust" (p. 143). Indeed, it is the idea of an elite cut-off from reality due to deceptive and self-deceptive groups of insulated "professional problem solvers" that formed one aspect of Hannah Arendt's (1973, p. 9) seminal commentary on *The Pentagon Papers*. These official documents, commissioned

by US Secretary of Defense Robert McNamara and leaked to the *New York Times* in 1971, revealed the disjuncture between the pessimistic intelligence assessments regarding the Vietnam War and official claims regarding both the course of the war and the reasons for US involvement (Ellsberg, 2003; Sheehan, 1971). For Arendt (1973, p. 12), whilst the raw intelligence reports were accurate, the professional problem solvers sought to erase inconvenient facts to such an extent that their assessment became detached from reality. Arendt (1973) concludes that because a US President is so reliant upon advisors as a source of information as to what is going on, he or she may become the most vulnerable to "complete manipulation" (p. 8).

Of course, beyond the question of rational and informed decision-making and harms to the body politic lies the concern that deception can become a powerful tool with which to exercise political power and engenders corruption and abuse. From a Marxian "false consciousness" perspective, ideology (Althusser, 1969), or its looser variant hegemony (Hall, 1977), can be seen as, at least in part, deceptive and a necessary, or at least persistent, part of ensuring that those in subservient political and economic positions either believe that their position in society is justified and appropriate, or at the very least acquiesce (Miller, 2002). Herman and Chomsky's (1988) propaganda model of the media presents a contemporary application of these ideas. Noting the close proximity of corporate media to political power, its profit orientation and propensity to relay official positions on specific policies as well as wider societal ideological frameworks, they show how mainstream US media present a profoundly distorted view of the world. As a result, crimes committed by official enemies are highlighted or even invented whilst those of allies, or "client states" as Herman and Chomsky (1988) call them, are ignored or downplayed. Most importantly, nefarious and immoral actions by the US government are systematically disguised by the corporate media. The net result is that varying proportions of US publics are led to perceive their own government as inherently law abiding, benign, and committed to high moral standards when the reality is very different. Herman and Chomsky (1988) document the US track record of supporting brutal regimes whilst deterring democracy in other states and prosecuting aggressive wars that are illegal under international law.

PROPAGANDA AND "PUBLIC RELATIONS"

So far we have discussed three broad strands of thinking with respect to the matter of deception, from those who see it as a necessary political strategy whether for ideological or practical reasons, those see it as having deleterious consequences for democracy, and Marxian perspectives for which deception is one important element with respect to how structures of inequality and exploitation are maintained. Questions with respect to how propaganda and deception are actually realised in practice are, however, poorly understood

by the aforementioned literatures. Deception is taken as a given whilst the organisations, institutions, doctrines, and practices that help to make deception a reality are effectively black boxed. It is here that two extant literatures, one on public relations (and related fields) and one on propaganda, provide at least a starting point for helping to understand deception as a political practice.

From Propaganda to Public Relations

Organised persuasive communication (OPC) (Bakir et al., 2018a) refers to intentional actions aimed at shaping both beliefs and conduct in order to achieve specific political outcomes. Historically, these activities have been labelled as propaganda, and in the first part of the twentieth century, leading thinkers such as political scientist Harold Lasswell (1927, 1935, 1951) and journalist Walter Lippman (1922, 1925, 1955) described the need for publics to be managed and manipulated in liberal democratic states. As Edward Bernays (1928) famously described:

> The conscious and intelligent manipulation of the organised habits and opinions of the masses is an important element in democratic society. Those who manipulate this unseen mechanism of society constitute an invisible government which is the true ruling power of our country. (p. 37)

The term "propaganda", however, came to be understood widely as a process of manipulation and one that frequently involved deception and new terms were invented. As Bernays explained, "propaganda got to be a bad word because of the Germans…using it [during WW1]. So what I did was to…find some other words. So we found the words Council on Public Relations".[1] Since the birth of public relations, or PR, a large variety of terms have come to be used to explain OPC activities including *political marketing*, *promotional culture*, *public diplomacy*, *strategic communication*, *perception management*, *political communication*, *public affairs*, *information operations*, *influence operations*, *political warfare*, and *advertising*. Although not necessarily always involving either manipulation or deception, these activities frequently do, and Philip Taylor (2002) has argued that they rightly be understood as euphemistic distractions:

> Let us first dispel with the euphemistic nonsense that surrounds this topic and which does in fact obscure what we are actually talking about – namely propaganda. …an entire euphemism industry has developed to deflect attention away from the realities of what they do, ranging from 'spin doctoring' and 'public affairs' at the political level to 'international information' and 'perception management' at the military level. …despite the euphemism game, democracies have grown ever more sophisticated at conducting propaganda, however labelled, which only they deny to be propaganda in the first place. (p. 20)

Whatever one's preferred term, these organised approaches to persuasion and influence play major roles in the exercise of power across political and economic realms. The scale of activities is huge: between 1979 and 1998, the PR consultancy industry in the UK mushroomed by a factor of 31 (11-fold increase in real terms) and this sector has "acted largely for business interests" (Miller & Dinan, 2000, pp. 10–14, 29). UK and US governments spend large sums on promotional activities. For example, according to a 2002 report by the UK Foreign and Commonwealth Office, it spent £340 million annually on public diplomacy operations in London (Miller, 2004, p. 80). The US federal government spent $16 billion on outside advertising and PR contractors between 2002 and 2012 (*Washington Times*, 2012).

Understanding the scale of the "organised persuasive communication" activity provides us, then, with a starting point for beginning to understand quite how extensive deceptive and propagandistic communication actually is in contemporary democracies. Put another way, we can start to comprehend the scale of deception that might be taking place and, critically, unpack the black boxing of these processes, which we see across the literature (discussed in the first section of this chapter). Before proceeding to discuss some of the institutional and strategic details vis-à-vis deception and propaganda, it is important to introduce a caveat. OPC should not be understood as necessarily deceptive or manipulative. As Bakir et al. (2018a) explain, it is possible to conceive of OPC that meets, at least in ideal terms, Habermasian notions of rational persuasion. This is why the term "propaganda" can and should be preserved for forms of communication that are manipulative and avoided as a blanket terms covering all OPC activities. Persuasion that avoids deception, incentivisation, and coercion can be seen as relatively consensual and democratic (Bakir et al., 2018a). And it may well be the case that a proportion of OPC activities fall into this category. However, as Bakir et al. (2018a) document, much of the literature on PR (and related fields such as strategic communication and public diplomacy) rarely if ever engages with manipulative forms of persuasion involving deception, coercion, and incentivisation whilst the existing literature on propaganda has a poorly developed conceptualisation of deception. Attention to coercion and incentivisation is important here: frequently neglected across both propaganda and PR literature studies, these aspects of manipulative OPC highlight how persuasion frequently operates in relation to physical, sociopolitical, and economic contexts where incentives and threats are part of persuasive communication activities (Bakir et al., 2018a, p. 2). To a very large extent, propaganda, understood as manipulative OPC, is something that is argued to go on in other nations (i.e., non-democratic states) or is relegated to history and wartime (e.g., World War I, World War II, and the Cold War). Because of this intellectual tunnel vision, it is at least likely that deception plays a far greater role in contemporary liberal democratic politics than is commonly understood. And it is to a discussion of such activities and the institutions and doctrines that underpin them that we now turn.

Forms of Deception

Deception, as a strategy, can undertake a number of forms. Lying, although most frequently associated with notions of propaganda, is actually comparatively rare. The political costs of being found out in a lie are frequently fatal in political terms and, as the German proverb states, "lies have short legs" (Friedrich, 1943, pp. 78–79; see also Ellul, 1965, pp. 53–57), meaning that they are vulnerable to being readily revealed. This said, lying as a political strategy does occur. For example, documented instances of lying include the Iran-Contra and Watergate scandals. With respect to the former, US officials lied with respect to their knowledge of a covert operation that involved the supply of weaponry to Iran, which was, in turn, used to funnel money to the Contras in Nicaragua who were seeking to overthrow the Sandinista government (Wroe, 1992). With the famous Watergate case, Nixon's claims that he was unaware of the cover-up of the burglary of Democratic National Congress offices were revealed as lies when taped conversations came to light (Sheehan, 1971). A more recent case involves the now infamous claims made by US officials in the run-up to the Iraq invasion in 2003. Here, Mearsheimer (2011) argues that US officials knowingly lied when they claimed that they knew for certain, based upon intelligence, that Iraq possesses weapons of mass destruction (WMDs).

More frequently, deception can occur through manipulation of information involving distortion, omission (Herring & Robinson, 2014), and misdirection (Bakir, 2013). Distortion involves manipulating information in order to exaggerate or downplay particular facts. For example, Herring and Robinson (2014) document the ways in which the UK government manipulated information through both exaggeration and omission in order to create the misleading impression of a current and serious WMD threat from Iraq. This was achieved, for example, by producing an intelligence-based dossier on Iraqi WMD that *omitted* countries understood to possess more advanced WMD programmes from the dossier (North Korea, Iran, Syria, and Libya) in order to help obscure the fact that Iraq was not a particularly serious or current problem (Herring & Robinson, 2014). *Distortion* of information occurred when it was decided to use a piece of raw intelligence, from a "source on trial", which was sub-sourced to a claim from another individual who had only promised hard evidence at a future date, in order to strengthen the entire dossier in a way that suggested it was known that Iraq was actively producing chemical and biological weapons (Herring & Robinson, 2014).

The Iraq case is also an example of *coercive deception* (Bakir et al., 2018a). Here, threat exaggeration is used in order to persuade audiences by deceiving them into experiencing psychological pressure or fear of physical harm. Mearsheimer (2011), in *Why Leaders Lie*, notes the frequency with which state politicians employ the tactic of scaremongering in order to mobilise populations in support of particular foreign policy objectives. The pattern of deception through distortion and omission was also reflected in the US and

manifested itself notably in sound bites such as "don't let the smoking gun be a mushroom cloud over New York", designed to invoke the nightmare of terrorist nuclear strike on a US city. Such messages were deceptive, in that they involved distortion via exaggeration of the available intelligence assessments; but they were also coercive because they were clearly intended to appeal to peoples' fears and anxieties over the possibility of a chemical, biological, or nuclear strike. A similar argument can be made about propaganda during the Cold War period, a large amount of which was based upon fearmongering with regard to the threat posed by the other side (Rawnsley, 1999). For example, the spectre of the Red Army invading Western Europe was part of the popular mindset amongst Western publics, even though the likelihood of Red Army tanks rolling across Western Europe was minuscule, if non-existent. Cold War insecurity and paranoia can only explain part of the official discourse that highlighted this unlikely scenario.

A further category is *deception through misdirection* (Bakir, 2013), which entails producing and disseminating true information but which is intended to direct public attention away from problematic issues. For example, Bakir (2013) analysed British and US management of the public revelation of the Bush administration's secret torture-intelligence policy and British complicity. This entailed instigating numerous investigations and inquiries across the US and UK. These were utilised to misdirect attention to a narrow part of policy failure (e.g., inadequate military training on how to handle detainees) and away from deeper issues. These deeper issues included the existence of a secret torture-intelligence policy, torture through "Enhanced Interrogation Techniques", the fact that the CIA was central to this policy, and the complicity of other nations.

A key point to take on board here is that deceptive communication involves much more than the telling of bare-faced lies: the processes are usually more subtle manipulations of information that, in the final analysis, can create significant deceptions. When one recognises that *distortion*, *omission*, and *misdirection* are key facets of deception, and also how widespread such activities are across the political realm, one starts to appreciate how important deception is to contemporary politics.

Sites of Production and Circulation

Propaganda and "distorted" communication are not just the preserve, of course, of governments, and in the contemporary world, we can readily identify a number of important production sites: think tanks, NGOs, academia, and intelligence agency networks.

For example, think tanks can be used as vehicles in order to generate information and, frequently, operate in ways which reflect the interests and agenda of their sponsors (Smith, 1993). Although not necessarily always part of contributing towards manipulated and propagandised representations

of particular issues, sometimes they are, such as the Henry Jackson Society (HJS), a think tank founded in 2005 and presented as bipartisan (Griffin, Aked, Miller, & Marusek, 2015). The HJS, funded by an array of undisclosed donors, has been active in "promoting a strongly pro-Israel agenda, organizing anti-Islam activities [and] advocating a transatlantic military and security regime" (Griffin et al., 2015, p. 74). Interestingly, and as revealed in a leaked document, HJS also planned coordinated activities aimed at discrediting Noam Chomsky via influencing mainstream media journalists (Sayeed, 2016). Clearly, shaping the information environment and, arguably, manipulating opinions (a.k.a. propaganda) has been a central objective of this think tank.

NGOs have also been implicated, on occasions, in the unintentional circulation of propagandistic information. For example, during the Libyan war in 2011, human rights-related claims against the Libyan government circulated prior to the intervention, including in an Amnesty International press briefing (2011). After the intervention, however, an Amnesty International official "could not corroborate allegations of mass human rights violations by Gaddafi regime troops" (House of Commons Foreign Affairs Select Committee, 2016, p. 15). In the case of the 2011–present war in Syria, the White Helmets group are presented as an independent organisation set up to save civilians. However, one government document indicates that the organisation has been funded as part of broader attempts to support "moderate opposition to provide services for their communities and to contest new space", and to empower "legitimate local governance structures to deliver services [and giving] credibility to the moderate opposition".[2] As such, the White Helmets would appear to be part of a broader US/UK regime change strategy which has supported the overthrow of the existing Syrian government. At the same time, the White Helmets have served an important public relations purpose by providing "an invaluable reporting and advocacy role" and "confidence to statements made by UK and other international leaders made in condemnation of Russian actions"[3] (Mason, 2017). Because the White Helmets only operate in areas held by opposition groups, they can only present a partial picture of events. The utility of this organisation, intentional or not, for propaganda purposes is clear. Indeed, a film about the White Helmets was even awarded an Oscar in 2016.[4]

Academia is not immune from propaganda activities and can itself become part of the broader propaganda apparatus. For example, Herring and Robinson (2003) argued that, to a large extent, the filters identified in the propaganda model as acting upon the media are also relevant to academia. Reliance upon grants, wishing to curry favour with official sources, as well as ideological imperatives, all mean that academia is far less free from the effects of power than is often assumed (See also Coser, 1965; Flaks, 1991; Mills, 1968). For example, Simpson's *Science of Coercion* (1994) draws upon a variety of sources, including Freedom of Information (FoI) releases, and carefully documents the relationship between the fledgling academic discipline of

communication science/studies and US psychological operations (psy ops). He highlights powerfully the interdependence between the academy and the US government and makes a powerful case that, in a very fundamental sense, communication science/studies is shaped, to this day, by the imperatives of political power.

Finally, the intelligence services are key producers and disseminators of propaganda in contemporary liberal democracies. For example, long before the now notorious intelligence-based WMD allegations made against Iraq during the run-up to the 2003 invasion of Iraq, British intelligence was involved in manipulating evidence in order to promote the impression that Iraq had an ongoing WMD programme. From 1991 onwards, the MI6 Operation Rockingham was involved in cherry-picking intelligence from the UN weapons inspections (set up after Gulf War 1) in order to, as a former UN weapons inspector put it, skew "UK intelligence about Iraqi WMD towards a preordained outcome that was more in line with British government policy that it was reflective of the truth" (Scott Ritter cited in Curtis, 2004, p. 47). Such activities were geared towards influencing the UN Security Council but also most likely designed to help maintain public support for the UK sanctions regime against Iraq. Operation Mass Appeal, initiated in the late 1990s, was precisely geared towards influencing public opinion by exaggerating the threat posed by Iraqi WMD. Finally, propaganda activities extend beyond attempts to influence publics via mainstream media and include popular culture propaganda. For example, Schou (2016) has documented the close involvement between the CIA and Hollywood. The relationships here range from mutual exploitation, through co-optation and on to more direct patterns of censorship. The overall net objective is to manipulate beliefs and attitudes in ways that are conducive to the interests of the US government.

It is important to note here that, when discussing matters of political deception, there is frequently the danger of either presenting such processes as overly "conspiratorial" or suggesting some kind of unified and monolithic propaganda machine. First and foremost, careful thought needs to be given, when researching propaganda and deception, to the possibility that actors are misperceiving the information they receive, involved in self-deception, or simply oblivious to the deceptive nature of the process that they are involved in (Herring & Robinson, 2014). Second, and more importantly, the preceding discussion with respect to the sites of production suggests a complex and fluid reality, whereby a variety of groups and organisations work towards common goals, in circumstance where interests are shared, by proactively seeking to "shape the information space". The propaganda, and potential deception that might accompany these attempts to "mobilise bias" (Schattschneider, 1960), emerges from the concerted activities of multiple groups. As Miller (2002) explains, when explaining the importance of understanding how ideology and power work:

> Rather than seek power in some mysterious unobservable process of ideological interpellation or articulation, or simply in understanding language, we must seek it in the actions of real people in the (would-be) secret (but sometimes discoverable) low conspiracies which are a continuous and inevitable part of capitalist rule; in censorship, spin, lobbying, public relations, marketing and advertising; in the institutions of 'disinformation and distraction' as Raymond Williams put it. (p. 252)

Following on from this point, it should also be understood that ideology, understood as an interest linked world view which can be internalised, believed, understood and acted upon by individuals, can be both augmented and reinforced by the OPC activities we are describing here. Indeed, as much as ideology might influence OPC activities, it is also the case that these activities are a key part of the process through which ideology is "created". The key point here is that the line between ideology and propaganda is continually changing and, perhaps, at times difficult to discern. So, whether someone involved in an OPC campaign is producing material is doing so because of their pre-existing ideology, or whether they are knowingly generating deceptive propaganda is something that needs careful consideration when examining case studies and processes. Indeed, it is precisely this interface between propaganda and ideology which makes the study of OPC, propaganda and deception of such pivotal importance in understanding how political power is exercised.

Deception then, involving manipulation via propagandistic OPC, is a key strategy through which political power is exercised in democracies. Via subtle processes involving distortion, omission, and misdirection, peoples' beliefs and behaviours can be shaped. Less subtle processes include deception through lying and coercive deception, and in addition, the interplay between ideology and propaganda needs to be acknowledged. All of these processes are certainly never always successful, and powerful actors are not omnipotent. However, the resources devoted to these activities, the range of sites across which propaganda production occurs, and variety of forms that deception can take should alert us to the importance of these manipulative processes in contemporary democracies and make clear their importance for understanding how political power is exercised.

Propaganda Here and Now: Deception in the Digital Era

As noted in the introduction, current political debate has been substantially shaped by the perceived crisis over "fake news". In large part emerging from the 2016 US presidential campaign between Hillary Clinton and Donald Trump, it is now widely believed that the phenomenon of "fake news" has reached epidemic proportions. For many, the problem of "fake news" is located either in the realm of social media, including alternative and independent media, or is a problem emanating from foreign sources. For example,

Bennett and Livingston (2018) argue that the problem is largely associated with "nationalist (primarily radical right) and foreign (commonly Russian) strategies to undermine institutional legitimacy and destabilise centre parties, governments and elections" (Bennett & Livingston, 2018, p. 122). This particular definition of the problem has been reflected by much of the emerging empirical research that is concerned with the spread of disinformation through social media and alternative/independent media. As much as this may be a real problem, "fake news", disinformation, and propaganda have also emanated from established political parties, vested interests, and mainstream media. For example, a highly significant political narrative of the last two years has concerned the Russia-gate scandal involving multiple allegations that Trump has colluded with Russia and that Russia actively interfered in the US election, in part through spreading "fake news". Despite all of the attention to the Russian collusion allegation, little substantive evidence appears to have emerged from the year-long Senate inquiry. Indeed, it has now transpired that the notorious Trump dossier detailing allegations of links with Russia, which appeared to have played a key role with respect to instigating investigations into President Trump, was in fact commissioned by the Democratic National Committee (DNC) and was authored by ex-UK MI6 intelligence officer Jonathan Steele. The issue is now subject to legal proceedings and there is one hypothesis that Russia-gate itself was a propaganda campaign designed to marginalise the Trump administration (McGovern, 2018; see also McKeigue, Miller, & Robinson, 2018). At the same time, there has been little sustained mainstream media attention to the content of the DNC leaks/hack which have fuelled so much of the controversy regarding the US elections and alleged Russian *information warfare*. Indeed, these leaked/hacked emails, released by WikiLeaks, indicate that the DNC actively favoured Clinton[5] over Bernie Sanders during the primaries whilst evidence of question fixing with CNN was also evidenced.[6] There are no serious challenges to the authenticity of these emails, and as such, they do not appear to be actual examples of "fake news". This did not, however, stop mainstream media linking Russia with the leaks and conflating all of this with a "fake news"/propaganda narrative.[7]

A similar pattern of omission and distortion can be seen with respect to the prevalent discourse that presents Russia as the new threat to the West where we would appear to be witnessing the emergence of a new "Cold War" and one that is focused on "hybrid" threats emanating from Russia. The general presentation of this problem focuses on Western attempts to counter a hostile and aggressive Russia with, for example, the creation of organisations designed to defend against Russian cyber warfare (European Commission, 2016, 2017). The central drift of the new "Cold War" discourse is that, beyond its alleged nefarious disinformation campaigns, Russia stands accused of aggression in the Ukraine and Syria. Whatever the accuracy of such claims, what is ignored in this particular narrative is Western support and involvement in multiple wars since 9/11 and transparently aggressive wars aimed at overthrowing "enemy" governments (specifically,

Afghanistan, Iraq, Libya, and Syria) (Robinson, 2017). Absent also from this narrative is the extensive involvement of the West in propaganda operations in recent years: for example, and with respect to the post 9/11 "war on terror", it was revealed by the UK government inquiry into the Iraq War that Tony Blair had expressed the need for a "tight knit propaganda" campaign (Robinson, 2017, p. 67), whilst the North Atlantic Treaty Organization (NATO) has sought to disseminate propaganda aimed at discrediting the Russian government. For example, NATO contracted the British firm Iota Global, to undertake anti Russian communications training. The firm was a part of the SCL group, which also included Cambridge Analytica (ironically, alleged to have played a role in Russian attempts to interfere in the US presidential election) (Cadwalladr & Graham-Harrison, 2018). According to a leaked document, confirmed as genuine by its author (Tatham, 2015a), this contract involved capacity building in anti-Russian strategic communications with Moldovan, Georgian, and Ukrainian participants (Tatham, 2015b). Another leaked document shows significant funding for Ukrainian organisations—including military and civil society groups—by the British government's "conflict pool" funding stream in 2014–2015. Amongst the objectives were to "discredit" Russian "political and military leaders".

Overall, then, it should be kept in mind that propaganda and deception are tools likely to be employed by a range of political actors, not simply those challenging establishment and mainstream positions within Western liberal democracies and "official enemies". Future research agendas should reflect this kind of objective and balanced approach, and avoid perceiving propaganda and deception as the preserve of only particular political actors. With this point in mind, what are some of the key digital propaganda and deception techniques which are now being employed?

Sock Puppets, Digital Propaganda, Bots, and Internet Censorship

In fact, from the point of view those engaged in deception and propaganda, the Internet environment facilitates a variety of invasive activities aimed at persuasion and influence. Many of these are manipulative. For instance, exploiting the digital age's capacity for "mass self-communication" (Castells, 2009) is the use of public relations techniques such as the "front group" where vested interests are disguised by ostensibly independent groups. Also, online identities can be assumed and used deceptively—a phenomena known as the "sock puppet"—a fake online persona. Although often used playfully, they are also used in economic and political influence strategies, such as Stella Artois (Watson, 2012) and the Special Operations Command of the US military (Fielding & Cobain, 2011). Another example of digital manipulation and deception concerns the work of Britain's Government Communications Headquarters (GCHQ), the Cheltenham-based communications and intelligence organization, revealed by the Snowden leaks. These show that the GCHQ's propaganda unit, the Joint Threat Research Intelligence Group (JTRIG), possesses a range of tools designed to alter the very fabric of digital

communication through online covert action including the publication of fake materials and deceptive content. For instance, "Clean sweep" is said to be able to "masquerade Facebook wall posts for individuals or entire countries"; "Gateway" can "artificially increase traffic to a website"; "Changeling" provides the "ability to spoof any email address and send email under that identity"; and "Havok" is a "real-time website cloning technique allowing on-the-fly alterations" (Greenwald, 2014).

Also, the rise of the political bot (Woolley & Howard, 2017) highlights the sophistication of strategies designed to exercise influence and control in the online environment. Social media bot technology involves "programs that communicate across multiple devices to perform some task" and "share the property of deploying messages and replicating themselves" (Woolley & Howard, 2017, p. 628). Whilst frequently used for activities relating to "spam, DDoS attacks, theft of confidential information", they can and have been used in overtly political contexts. As Woolley and Howard (2017) explain:

> In Mexico, bots have been used on Twitter by both ruling and minority parties. …bots are programmed to co-opt the opposition's hashtags and send out thousands of garbled or propaganda-laden tweets to block any counter-organizational or communication efforts. In the United States, the United Kingdom, and Australia bots have been used to pad politicians' follower lists. These fake followers can be purchased for nominal prices with the intent of making a user seem more popular or influential. (p. 630)

During the 2010 US midterm elections and Massachusetts special election, social bots were reportedly used to support some candidates and smear their opponents, by creating thousands of tweets pointing to websites with "fake news" (Metaxas & Mustafaraj, 2012).

Finally, at present there appears to be a concerted drive to develop methods of identifying so-called fake news with major Internet giants such as Facebook, Google, and Twitter moving towards a substantial attempt to manipulate information on the Internet so as to limit information claimed or perceived to be "fake news". Some of this might be being driven by pressure from governments, and any resulting use of artificial intelligence (AI) in order to filter information perceived to be bogus may lead to a significant degree of censorship across the Internet. The very real danger here is that these developments lead to a situation where the Internet loses its open and democratic potential and becomes, broadly speaking, a tool for manipulation by powerful actors in society.

Conclusion

Deception and propaganda are, then, alive and well in contemporary politics. As a political strategy rationalised and justified by a number of ideological positions (e.g., realism and neoconservatism amongst others), its use is

prevalent not only by the "usual suspects", such as authoritarian or totalitarian states, but also within contemporary liberal democracies. As a practice, deceptive propaganda is inherently undemocratic, although under certain exceptional circumstances it can be justified, and can contribute to a significant erosion of the public sphere and democratic politics. Indeed, the scale of propaganda activities, the resources devoted to them, and range of sites across which such activities can be identified (e.g., think tanks, NGOs, academia, intelligence services), coupled with the relatively discrete ways through which deception can be realised (e.g., omission, distortion, misdirection), means that it is highly likely that many social scientists are undermeasuring the scale of these activities and, accordingly, the problems they may be raising for democracy. Add to this the opportunities now provided by the Internet and digital communication for manipulation of information, it should be clear that there is a powerful case for sustained attention to deception and propaganda.

Future research should be directed as follows: first, case studies exploring deception across the full range of political activity are necessary to establish the scale and extent of deception in the contemporary political realm. Second, understanding the complex networks of actors involved in "shaping the information" environment, and their interests and objectives, can help social scientists understand and explain how information comes to be manipulated and the precise mechanisms that lie behind deceptive propaganda campaigns. Third, theoretical work is needed in order to understand better the interplay between ideology and deceptive propaganda. Fourth, normative political theory can be engaged in order to develop robust and fine-grained frameworks that can be employed in order to establish the exceptional circumstances when deceptive propaganda might be argued to be necessary and justified; also necessary here is the development of more ethical forms of persuasion which are both consensual and democratic (Bakir et al., 2018a). Fifth, digital and Internet-based propaganda needs significant attention so that we can better understand the tactics and tools now being employed by powerful actors. Finally, attention needs to be paid to developing educational strategies that provide citizens with the skills and knowledge to defend against deceptive propaganda and navigate successfully the diverse information sources now available.

Ultimately, there is much at stake. Deception as a political strategy and the ability of powerful actors to manipulate in the current digital environment mean that research and understanding are urgently needed into this issue area. And, for all the allegations and fearmongering, the biggest problem lies, arguably, not in the behaviour of some foreign actor (such as Russia) or independent/alternative media, but within existing institutions including mainstream media, corporate Internet giants, and democratic governments themselves. It is here that scholarly attention should be focused and with the primary objective being to defend the public sphere and democracy from practices of deception.

Notes

1. Bernays interview, 'Happiness Machines', *The Century of the Self*, Part 1, BBC2, April 29, 2002.
2. UK Gov summary document available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/630409/Syria_Resilience_2017.pdf.
3. Ibid.
4. The White Helmets and other matters relating to UK government 'info ops' in Syria are currently being researched by some members of the Working Group on Syria, Propaganda and Media http://syriapropagandamedia.org.
5. For example, see https://www.washingtonpost.com/news/the-fix/wp/2016/07/24/here-are-the-latest-most-damaging-things-in-the-dncs-leaked-emails/?utm_term=.52fe1cfcdd6c.
6. For example, see https://www.washingtonpost.com/lifestyle/style/cnn-drops-donna-brazile-as-pundit-over-wikileaks-revelations/2016/10/31/2f1c6abc-9f92-11e6-8d63-3e0a660f1f04_story.html?utm_term=.9f71d28bf2d3.
7. For example, see https://www.washingtonpost.com/business/economy/russian-propaganda-effort-helped-spread-fake-news-during-election-experts-say/2016/11/24/793903b6-8a40-4ca9-b712-716af66098fe_story.html?utm_term=.ab301a2365a0.

References

Althusser, L. (1969). *For Marx*. London: Allen Lane.

Amnesty International. (2011). Libya: Organization calls for immediate arms embargo and assets freeze. Retrieved from https://www.amnestyusa.org/press-releases/libya-organization-calls-for-immediate-arms-embargo-and-assets-freeze/.

Arendt, H. (1973). Lying in politics. In *Crises of the Republic*. San Diego, CA: Harcourt, Brace & Co.

Bakir, V. (2013). *Torture, intelligence and sousveillance in the war on terror: Agenda-building struggles*. Farnham: Ashgate.

Bakir, V., Herring. E., Miller, D., & Robinson, P. (2018a). Organised persuasive communication: A new conceptual framework for research on public relations, propaganda and promotional culture. *Critical Sociology*. Available online first.

Bakir, V., Herring, E., Miller, D., & Robinson, P. (2018b). Deception and lying in politics. In Jorg Mauber (Ed.), *Oxford handbook on lying*. Oxford: Oxford University Press.

Bennett, W. L., & Livingston, S. L. (2018). The disinformation disorder: Disruptive communication and the decline of democratic institutions. *European Journal of Communication, 33*(2), 122–139.

Bernays, E. (1928). *Propaganda*. New York: Horace Liveright.

Bok, S. (1999). *Lying: Moral choice in public and private life*. Sussex: Harvester.

Cadwalladr, C., & Graham-Harrison, E. (2018). Cambridge Analytica: Links to Moscow oil firm and St Petersburg University. *The Guardian*. Retrieved from https://www.theguardian.com/news/2018/mar/17/cambridge-academic-trawling-facebook-had-links-to-russian-university.

Castells, M. (2009). *Communication power*. Oxford: Oxford University Press.

Cliffe, L., Ramsay, M., & Bartlett, D. (2000). *The politics of lying: Implications for democracy.* Basingstoke: Macmillan.

Corner, J. (2007). Mediated politics, promotional culture and the idea of propaganda. *Media, Culture and Society, 29*(4), 669–677.

Coser, L. (1965). *Men of ideas.* New York: Simon and Schuster.

Curtis, M. (2004). *Unpeople: Britain's secret human rights abuses.* London: Vintage.

Dallek, R. (1979). *Franklin D. Roosevelt and American foreign policy 1932–1945.* New York: Oxford University Press.

Dentith, M. (2018). The problem of conspiracism. *Argumenta, 3*(2), 327–343.

Dick, F. (1991). Making history and making theory: Notes on how intellectuals seek relevance. In Charles Lemert (Ed.), *Intellectuals and politics: Social theory in a changing world.* Newbury Park, CA: Sage.

Ellsberg, D. (2003). *Secrets: A memoir of Vietnam and the Pentagon papers.* London: Penguin.

Ellul, J. (1965). *Propaganda: The formation of men's attitudes.* New York: Knopf.

European Commission. (2016). *Joint communication to the European Parliament and the council: Joint framework on countering hybrid threats—A European Union response.* Brussels, 6.4.2016. JOIN(2016) 18 final. Available at https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52016JC0018.

European Commission. (2017). EU welcomes establishment of the Finnish centre of excellence for countering hybrid threats. Retrieved from https://eeas.europa.eu/headquarters/headQuarters-homepage/24572/eu-welcomes-establishment-finnish-centre-excellence-countering-hybrid-threats_en.

Fielding, N., & Cobain, I. (2011). Revealed: US spy operation that manipulates social media. *The Guardian.* Retrieved from https://www.theguardian.com/technology/2011/mar/17/us-spy-operation-social-networks.

Friedrich, C. J. (1943). Issues of informational strategy. *Public Opinion Quarterly, 7*(1), 77–89.

Greenwald, G. (2014). How covert agents infiltrate the internet to manipulate and deceive. *The Intercept.* Retrieved from https://theintercept.com/2014/02/24/jtrig-manipulation/.

Greenwald, G., & Fishman, A. (2015). Controversial GCHQ unit engaged in domestic law enforcement, online propaganda, psychology research. *The Intercept.* Retrieved from https://theintercept.com/2015/06/22/controversial-gchq-unit-domestic-law-enforcement-propaganda/.

Griffin, T., Aked, H., Miller, D., & Marusek, S. (2015). *The Henry Jackson society and the degeneration of British neoconservatism: Liberal interventionism, islamophobia and the 'war on terror'.* Glasgow: Public Interest Investigations. Retrieved from http://www.spinwatch.org/index.php/issues/more/item/5777-new-report-on-the-henry-jackson-society.

Habermas, J. (1984). *The theory of communicative action, volume 1: Reason and the rationalization of society* (T. McCarthy, Trans.). Boston, MA: Beacon Press.

Hall, S. (1977). Culture, the media and the ideological effect. In J. Curran, M. Gurevitch, & J. Woollacott (Eds.), *Mass communication and society.* London: Edward Arnold.

Herman, N., & Chomsky, N. (1988). *Manufacturing consent: The political economy of the mass media.* New York: Pantheon.

Herring, E., & Robinson, P. (2003). Too polemical or too critical? Chomsky on the study of the news media and US foreign policy. *Review of International Studies, 29*(4), 553–568.

Herring, E., & Robinson, P. (2014). Report X marks the spot: The British Government's deceptive dossier on Iraq and WMD. *Political Science Quarterly, 129*(4), 551–584.

House of Commons Foreign Affairs Committee. (2016). *Libya: Examination of intervention and collapse and the UK's future policy options.* London: HMSO.

Lasswell, H. (1927). *Propaganda technique in the World War.* Cambridge: MIT Press.

Lasswell, H. (1951). Political and psychological warfare. In Daniel Lerner (Ed.), *Propaganda in war and crisis* (pp. 261–266). New York: George Stewart.

Lasswell, H., Casey, R., & Smith, B. L. (1935). *Propaganda and promotional activities: An annotated bibliography.* Chicago: University of Chicago Press.

Lebow, R. N. (2008). The Ancient Greeks and modern realism: Ethics, persuasion, and power. *Philosophy and Social Science, 9*(1), 2–16.

Lippman, W. (1922). *Public opinion.* New York: Harcourt, Brace.

Lippman, W. (1925). *The phantom public.* New York: Harcourt, Brace.

Lippman, W. (1955). *Essays in the public philosophy.* Boston, MA: Little, Brown.

Machiavelli, N. (2003 [1532]). *The prince* (George Bull, Trans.). London: Penguin.

Mason, J. (2017). Power, petroleum and propaganda: Identifying the role of the White Helmets in the Syrian Civil War (MA Dissertation, University of Sheffield).

McGovern, R. (2018). The FBI hand behind Russia-gate. *Consortiumnews.* Available at https://consortiumnews.com/2018/01/11/the-fbi-hand-behind-russia-gate/.

McKeigue, P., Miller, D., & Robinson, P. (2018). Briefing note: Update on the Salisbury poisonings. *Working Group on Syria, Propaganda and Media.* Retrieved from http://syriapropagandamedia.org/working-papers/briefing-note-update-on-the-salisbury-poisonings.

Mearsheimer, J. (2011). *Why leaders lie: The truth about lying in international politics.* New York: Oxford University Press; London: Duckworth Overlook.

Metaxas, P. T., & Mustafaraj, E. (2012). Social media and the elections. *Science, 338*(6106), 472–473.

Miller, D. (2002). Media power and class power: Overplaying ideology. In L. Panitch, & C. Leys (Eds.), *'A World of Contradictions', Socialist Register 2001.* London: Merlin Press; New York: Monthly Review Press.

Miller, D. (2004). The propaganda machine. In D. Miller (Ed.), *Tell me lies: Propaganda and media distortion in the attack on Iraq* (pp. 80–99). London: Pluto Press.

Miller, D., Brown, L., Dinan, W., & Stavinoha, L. (Eds.). (in preparation). *Researching the powerful: Public sociology in action.* London: Routledge.

Miller, D., & Dinan, W. (2000). The Rise of the PR industry in Britain 1979–1998. *European Journal of Communication, 15*(1), 5–35.

Mills, C. W. (1968). *Power, politics and people: Collected essays.* Oxford: Oxford University Press.

Oborne, P. (2014). *The rise of political lying.* London: Simon and Schuster.

Plato. (360 BC). *The republic* (B. Jowett, Trans.), The internet classics archive. Retrieved from http://classics.mit.edu/Plato/republic.html.

Ramsay, M. (2000). Justifications for lying in politics. In L. Cliffe, M. Ramsay, & D. Bartlett (Eds.), *The politics of lying* (pp. 3–26). London: Macmillan.

Robinson, P. (2017a). Learning from the Chilcot report: Propaganda, deception and the "war on terror". *International Journal of Contemporary Iraqi Studies, 11*(1&2), 47–73.

Robinson, P. (2017b). The CNN effect and humanitarian intervention. In P. Robinson, R. Froehlich, & P. Seib (Eds.), *The Routledge handbook of media, conflict and Syria* (pp. 91–105). London: Routledge.

Robinson, P. (2018). Extending the propaganda model: Analysing propaganda strategies and sites of production. In J. Pedro-Caranana, D. Broudy, & J. Klaehn (Eds.), *The propaganda model today: Debating herman and Chomsky.* London: University of Westminster Press.

Sayeed, T. (2016). Chomsky and his critics. *Mondoweiss.* February 19, 2016. http://mondoweiss.net/2016/02/chomsky-and-his-critics/.

Schattschneider, E. E. (1960). *The semi-sovereign people: A realist's view of democracy in America.* New York: Holt, Rinehart and Winston.

Schou, N. (2016). *Spooked: How the CIA manipulates the media and hoodwinks hollywood.* New York: Skyhorse Publishing.

Sheehan, N. (1971). *The Pentagon Papers.* New York, NY: Bantam.

Simpson, C. (1994). *The Science of coercion: Communication research and psychological warfare, 1945–1960.* Oxford: Oxford University Press.

Smith, J. A. (1993). *Idea brokers: Think tanks and the rise of the new policy elite.* New York: The Free Press.

Strauss, L. (1958). *Thoughts on Machiavelli.* Chicago, IL: University of Chicago Press.

Strauss, L. (1975). *The argument and the action of Plato's laws.* Chicago: IL University of Chicago Press.

Tatham, S. (2015a). The solution to Russian propaganda is not EU or NATO propaganda but advanced social science to understand and mitigate its effect in targeted populations. Policy Paper, no. 4, July. Riga, Latvia: Center for Security and Strategic Research, National Defence Academy of Latvia. https://www.stratcom-coe.org/steve-tatham-solution-russian-propaganda-not-eu-or-nato-propaganda-advanced-social-science.

Tatham, S. (2015b). Coordination meeting: On StratCom training for Ukraine, Georgia and Moldova. https://drakulablogdotcom3.files.wordpress.com/2015/04/stratcom-training.doc.

Taylor, P. M. (2002). Perception management and the 'war' against terrorism. *Journal of Information Warfare, 1*(3), 16–29.

Washington Times. (2012). Feds spent $16B since '02 on outside PR, ads. Retrieved from https://www.washingtontimes.com/news/2012/nov/25/feds-spent-16b-since-02-on-outside-pr-ads/.

Watson, T. (2012). The lobbyists, the Russians, Google and 'wife beater'. *Socialist Unity.* Retrieved from https://socialistunity.com/the-lobbyists-the-russians-google-and-wife-beater/.

Woolley, S. C., & Howard, P. N. (2017). Computational propaganda worldwide: Executive summary. *Computational Propaganda Research Project.* Retrieved from http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf.

Wroe, A. (1992). *Lives, lies and the Iran-Contra affair.* London: I.B. Tauris.

CHAPTER 51

Propaganda, Survival, and Living to Tell the Truth: An Analysis of North Korean Refugee Memoirs

Tony Docan-Morgan, Sarah A. Son and Golnar B. Teimouri

Recent international events have heightened global interest in the North Korean regime's engagement with the world; however, information about North Korea, known formally as the Democratic People's Republic of Korea, remains largely nebulous, due to the continuing secrecy of the regime. Although governmental and non-governmental organizations, journalists, and scholars observing North Korea from the outside have pointed to the regime's use of ruthlessness and deception, we know little about the ground realities of communication and specifically how strategies of deception and truth are perceived and used by individuals in North Korea. The current chapter utilizes North Korean refugee narratives to examine these issues in close detail. Doing so better informs us about the regime's ubiquitous strategies of human control, works toward highlighting the voices of those who have lived under oppressive conditions, and advocates for further attention and action that must be directed to issues of human rights.

T. Docan-Morgan (✉)
Department of Communication Studies, University of Wisconsin-La Crosse,
La Crosse, WI, USA
e-mail: tdocan@uwlax.edu

S. A. Son
Korea University Human Rights Center, Seoul, South Korea

G. B. Teimouri
University of Chicago, Chicago, USA

We begin by providing a brief historical context before describing the conditions in North Korea that precipitated the two-decade-long refugee crisis that emerged around the time of the *Great Famine* or *Arduous March*, and which has remained largely beyond the reach of the international community due to complex regional geopolitical factors. We then discuss the impact of the long and perilous journey from North to South in terms of identity formation and the consequent need to recount the story of that journey in different settings, including in the form of memoirs for public dissemination. Following on from this, we provide a thematic analysis of North Korean refugee accounts and call for purposeful research and action pertaining to North Korean human rights and North Korean refugee adjustment and advancement.

BACKGROUND: A NATION DIVIDED AND OPPRESSED

Few geographical spaces have experienced as much turmoil, dispossession, conflict, and tragedy as the Korean Peninsula during the twentieth century. Although South Korea has seen remarkable achievements in the form of its rapid economic development and subsequent democratization, North Korea continues to maintain a reputation as the most closed state on earth, as well as the world's worst violator of human rights. How did a region that enjoyed millennia of political unity under successive domestic royal dynasties reach this point? Numerous compelling accounts of twentieth-century Korean history document in detail the events of this period, emphasizing the damage done as a consequence of Japan's 35-year colonial occupation of the Peninsula beginning in 1910 (Cumings, 2005; Oberdorfer, 2001; Shin, 2018). Japan's heavy-handed rule led to a Korean resistance movement that was divided in ideological orientation (Robinson, 1982). When the Japanese surrendered to the allies in August 1945, the US and the Soviet Union found themselves in a contest for control of the Peninsula, resulting in the hurried division of North and South. Officers in Washington, "…[w]orking in haste and under great pressure, and using a *National Geographic* map for reference…proposed that U.S. troops occupy the area south of the thirty-eighth parallel…and that Soviet troops occupy the area north of the parallel" (Oberdorfer, 2001, p. 6). It was intended that this situation would only be temporary; however, the plan to unite the two regions never came to fruition. Continuous negotiations failed, resulting in the eventual conversion of the former colonial resistance movements into the formation of two separate governments by 1948—one built on a socialist model backed by the Soviets and the Chinese in the North, and a US-backed capitalist regime in the South (Hassig & Oh, 2015; Lankov, 2013).

The North Korean leader, Kim Il-Sung[1] (grandfather of the current leader, Kim Jong-Un), had a stated goal to extend his control over the southern half of the peninsula (Hassig & Oh, 2015, p. 16). And so his forces invaded

South Korea in June 1950, initiating a devastating three-year war that ended where it started, roughly along the 38th Parallel—today the location of the Demilitarized Zone (DMZ). The DMZ is often referred to as one of the most heavily guarded areas in the world, and no unauthorized movement is permitted across it. Although both Koreas began an earnest effort to rebuild in the years following the war according to their respective economic models, the end of the Cold War and collapse of the Soviet bloc had a grave impact on North Korea economically. Without its former communist allies to continue contributing economic support, the North Korean public distribution system (PDS) began to collapse, having been in gradual decline for some time (Lankov, 2013, pp. 78–79). By the early 1990s, bad weather and poor harvests had drained almost all resources from the system, and widespread famine ravaged the country, hitting rural areas especially hard. North Korean defectors to South Korea have told of schools emptying, factories closing and falling into disrepair, blackouts becoming common, and health care declining into obscurity. While starving citizens foraged for wild plants and small animals to feed themselves, thousands of others began to flee across the northern border to China in search of a way out of their misery (Demick, 2010; Lankov, 2013, pp. 88–90). With them, these famine refugees brought the first real accounts of the extraordinary brutality of a regime which had maintained a system of extraordinary social control to ensure its survival, even while its people starved. Twenty-five years on, the regime continues to exert many of the same methods of control.

The North Korean regime, according to Kirkpatrick (2014), fears two things:

> *the outflow of its citizens* and *the inflow of information*. Pyongyang's crackdown on citizens who try to leave reflects the essential insecurity at the core of every totalitarian regime. So, too, does its suppression of information coming from any source other than itself. It is the response of the government that understands just how subversive the truth can be if a significant segment of its population is exposed to it. The regime knows that information, if spread, threatens the very essence of its power. This gives it a powerful incentive to keep its citizens from encountering any and all unauthorized information. (p. 298; italics added)

The need to isolate its people from knowledge and understanding of the outside world has led to abuses of human rights on a scale that places North Korea consistently at the top of the list of countries worldwide engaging in systematic rights violations. In 2014, The United Nations Commission of Inquiry on Human Rights in the Democratic People's Republic of Korea (UN COI) reported the following:

> Systematic, widespread and gross human rights violations have been and are being committed by the Democratic People's Republic of Korea, its institutions and officials. In many instances, the violations of human rights found by the

commission constitute crimes against humanity. These are not mere excesses of the State; they are essential components of a political system that has moved far from the ideals on which it claims to be founded. The gravity, scale and nature of these violations reveal a State that does not have any parallel in the contemporary world. Political scientists of the twentieth century characterized this type of political organization as a totalitarian State: a State that does not content itself with ensuring the authoritarian rule of a small group of people, but seeks to dominate every aspect of its citizens' lives and terrorizes them from within. (p. 365)

The report also addresses the DPRK's indoctrination of citizens from childhood, "suppressing all political and religious expression that questions the official ideology, and tightly controlling citizens' physical movement and their means of communication with each other and with those in other countries" (p. 365). Regarding the right to food, "The State's monopolization of access to food has been used as an important means to enforce political loyalty. The distribution of food has prioritized those who are useful to the survival of the current political system at the expense of those deemed to be expendable. Citizens' complete dependence on the State led to one of the worst cases of famine in recent history" (p. 366). The famine of the mid-1990s killed up to one million people according to some reports (Haggard & Noland, 2011).[2] Further, the COI found that "The keystone to the political system is the vast political and security apparatus that strategically uses surveillance, coercion, fear and punishment to preclude the expression of any dissent. Public executions and enforced disappearance to political prison camps serve as the ultimate means to terrorize the population into submission" (p. 366). The UN COI was by no means the first report to expose such facts: Human rights groups in South Korea and internationally have been investigating and publishing such data for over two decades. However, the COI report's conclusion that North Korea's behavior constitutes crimes against humanity in international law, and its consequent recommendation that North Korea's leadership, including Kim Jong-Un, be referred to the International Criminal Court, pointed to the seriousness of the situation under which North Koreans continue to suffer (Son, 2018).

Escaping North Korea

Given the grave conditions described above, it is not surprising that tens of thousands of North Koreans have attempted to escape the country, primarily since the mid- to late-1990s when the *Great Famine* was biting in earnest. Prior to that time, only a handful of North Koreans had left and sought refuge elsewhere. Refugees, defined as "people who are forced to flee for safety reasons from their country of origin due to war, fear of persecution, or famine" (Sorrells, 2012, p. 132), report various reasons for leaving North Korea. Although responses differ depending on factors such as when one defected, motivations have included hunger and the search for food (Fahy, 2015), loss

of status, frustration over lack of opportunities, political persecution due to one's family history, a desire to live in better conditions (Human Rights Watch, 2002), following another who had already left the country (Lee, 2006), economic conditions, lack of political and religious freedom, and fear (Fahy, 2015; Haggard & Noland, 2011). Although it is impossible to determine accurately, various sources place the number of North Korean refugees at around 100,000 (e.g., Haggard & Noland, 2011) or higher (Kim, 2008), including those hiding in China,[3] as well as those who have found refuge in countries that have allowed them to apply for asylum or, in South Korea's case, given them citizenship. Although accurate demographic information on the refugee population in China is hard to come by, the vast majority of those who reach South Korea are women. Of the 32,467 total arrivals to South Korea (as projected to the end of 2018), 72% are women, and that proportion is rising (Ministry of Unification, 2018).[4] High-level defections also occur; loyalists, including upper class and senior officials, may defect because of purge fears (which can include prison, torture, and death), being accused of crimes, reduced political enthusiasm, and concern over national prospects (Lankov, 2017). These high-level defections can happen abroad in the case of diplomats or overseas workers, as was the case with Thae Yong Ho—a high-level diplomat who defected with his family from the North Korean embassy in London in 2016.

But how do ordinary North Korean refugees make it out of North Korea without getting caught and repatriated? Nearly all of those who attempt to escape do so via the Sino-North Korean border, which spans 1420 kilometers (880 miles). Haggard and Noland's (2011) survey of North Korean refugees in China and South Korea provides insight on the "mechanics of escape." They found that three-quarters of respondents in their China survey indicated receiving help, and half of respondents indicated paying for assistance. Haggard and Noland (2011) suggest "bribery of officials and/or the emergence of a group of brokers or 'coyotes' plays a large role in their escape" (p. 32). Other responses point to family or friends and non-governmental organizations aiding in the escapes. Although periodic crackdowns by both the Chinese and North Korean authorities can alter the patterns and means of escape, the primary networks that ultimately lead to safe passage continue to involve key players including "governments, missionaries, brokers and diplomatic missions from Ulaan Bataar to Rangoon" (International Crisis Group, 2006, p. 1), as well as NGOs such as Liberty in North Korea, which fundraises internationally to support safer passage of refugees "without cost or condition" (Liberty in North Korea, n.d.). Fieldwork studies, often involving observation and interviews, corroborate and extend these findings (see International Crisis Group, 2006; Kirkpatrick, 2014). However, this body of work has not yet examined what specific communicative strategies refugees report using as they flee North Korea, live on the margins in China, and ultimately attempt to exit in order to permanently resettle elsewhere.

As indicated previously, leaving North Korea without permission is considered a serious crime, with passage through China presenting severe obstacles. The 1986 Border Area Affairs Agreement and the 1960 Escaped Criminals Extradition Treaty between North Korea and China enable repatriation of any North Koreans caught in Chinese territory (International Crisis Group, 2006, p. 2), where they are branded "illegal economic migrants" (Committee for Human Rights in North Korea, 2009). Those who are caught and accused of defecting or attempting to receive asylum in a third country have faced lengthy prison sentences, forced labor, torture, confiscation of property, or death (Haggard & Noland, 2011). Sungju Lee (2016) sums up the issue:

> China views North Korean defectors not as refugees fleeing severe human rights abuses but as illegal work migrants. North Koreans caught in China are still deported to Joseon [North Korea], where they are imprisoned. North Koreans in China live perilous lives, flirting in an underground work economy and suffering abuse, poverty, and depraved living conditions. All they want is freedom. (p. 307)

North Koreans entering China do not necessarily have immediate plans to try to reach South Korea. Some end up in precarity while trading across the North Korea–China border, finding themselves unable or unwilling to return. Others are unsure where to move to next, and it can be a matter of sheer chance as to if and how they proceed onwards. Once in China, they often face three perilous paths—attempting to return to North Korea, surviving on the margins in China, or attempting to exit China via Mongolia or Southeast Asia to permanently resettle in South Korea or other countries that will accept them. While remaining in China, North Koreans are still highly vulnerable to continued human rights abuses, including imprisonment, being sold into prostitution or forced marriages with Chinese men, or forced to work in undesirable or dangerous conditions for little to no pay (International Crisis Group, 2006). If caught and forced back to North Korea, refugees are interrogated, labeled criminals and traitors, and can receive punishment in the form of incarceration in labor camps or the death penalty (Committee for Human Rights in North Korea, 2009). North Korean refugees are therefore in "an extraordinarily vulnerable population, and their current status and future prospects constitute a first-order humanitarian problem for the international community" (Haggard & Noland, 2011, pp. 1, 3).

The analysis in this chapter helps us better understand the experiences of North Korean defectors—all of whom have had their voices silenced while in North Korea and often thereafter on the run or even while adapting to their new place of legal citizenship. This analysis also provides voice for the realities, challenges, losses, and advances that North Korean defectors have experienced, and brings awareness to the many human rights violations that have occurred and continue to occur in North Korea. Speaking about North

Korean refugees, Kim (2008) urges, "It is also necessary for society to recognize and understand the testimonies of trauma victims" (p. 94).

North Korean Refugees/Defectors and Identity

The extraordinary and often traumatic experiences conferred upon North Korean refugees during their journeys are certainly life-altering, but the experience of turmoil does not necessarily end on arrival in a country where they are granted protection. It is at this point that a new struggle begins: finding a new identity in a new society. This is particularly true for those who end up in South Korea, given the highly politicized nature of their presence there. On arrival in South Korea, defectors are questioned at length by intelligence services to ensure their origin and the accuracy of their story, before spending several months in a closed facility (*Hanawon*) for debriefing, counseling, necessary medical treatment, and training in life skills for release into society. North Koreans in South Korea often comment on the struggles they face integrating into what is a highly homogenous society, due to differences in spoken accent and vocabulary, style, and even their physical features on account of the malnutrition many have experienced. Many suffer from post-traumatic stress disorder, depression and anxiety, and physical disabilities from injuries or poor medical care. Many come with a lack of marketable education and find themselves underemployed and suffering from financial hardship as a consequence. They also experience discrimination in the job market, despite the support provided by the government's defector settlement system (see Chung, 2008; Chung & Seo, 2007; International Crisis Group, 2011; Suh, 2013). All of this can lead to a sense of "identity crisis," whereby they feel that the North Korean aspect of their identity is inferior, or only of worth in certain, externally prescribed settings such as when invited to give testimonies by government ministries, NGOs, or churches—each of which may be using defector testimony for a certain instrumental end (International Crisis Group, 2011; Son, 2016). This experience is not unique to those who end up in South Korea: Once in America, Joseph Kim (2015) writes about being "desperate to fit in," feeling "like an alien" (p. 260), and the adjustment progress being "painful and slow" (p. 263). Multiple refugees speak of experiencing "survivors guilt" (e.g., J. Kim, 2015, p. 265). The experience of identity crisis is remarked upon by a number of the memoirists included in this study, and their comments also provide first-person insight regarding the perennially contentious issue of terminology regarding defectors/refugees from North Korea.[5]

When briefly discussing her identity in her introduction to her memoir, Hyeonseo Lee (2015) states, "I have come to accept that as a North Korean defector I am an outsider in the world. An exile" (p. xii). Further, she states, "The simple solution to my problem of identity is to say I am Korean, but there is no such nation. The single Korea does not exist" (p. xiii). Later, she

also uses the phrase, "North Koreans in the South" (p. 281). Similarly, Sungju Lee (2017) has publically discussed his identity and settled on the term "Korean." In his memoir, Sungju Lee (2016) refers to himself as "a North Korean defector with South Korean citizenship" (p. 300) and generally uses the terms "North Korean defectors" or simply "defectors." In her memoir, Yeonmi Park (2015) mentions, "Many of us who have escaped call ourselves 'defectors' because by refusing to accept our fate and die for the Leader, we have deserted our duty. The regime calls us traitors. If I tried to return, I would be executed" (p. 4). In his memoir, Chol-hwan Kang (2001) uses the terms "North Korean renegade" and "refugee" (pp. 223–233). Joseph Kim (2015) uses the term "refugee" when referring to himself and other North Koreans on the run in China (p. 230) and once he finds freedom in the US (p. 261).

Eunsun Kim (2015) uses various terms: "North Korean escapees" (pp. 125, 199), "defectors from North Korea" (p. 160), "defectors" (p. 178), and "my fellow North Koreans" (p. 213). Yet later, once settled in South Korea with her mother she states, "We were now 'normal' South Koreans" (p. 182), and when planning a trip to China to visit Eunsun Kim's sister, she mentions, "It was possible now because we were true South Koreans, and no longer illegal migrants" (p. 191). At the end of her book, Eunsun Kim (2015) states, "As for me, I simply consider myself, above all, Korean" (p. 224). These many labels speak to the multi-faceted and complicated identities of individuals who have left the North and resettled elsewhere.[6] Perhaps more important is that individuals who permanently leave North Korea must deal with identity politics surrounding their status in their new society, where it can be difficult to find a voice to represent their experience. This is particularly true in South Korea where the political divide over North Korea runs deep, and where North Korean defectors are often positioned as a litmus test of a long-hoped-for inter-Korean unification process, an expectation which some find burdensome (Son, 2015). This perhaps lends to the appeal of writing and publishing a memoir, especially if done so in both Korean and English, to allow the memoirist to present a version of events that he or she has crafted independently of any state or official discourse, as well as a measure of detailed story-telling impossible in most other situations or forms of media.[7] Further, "narratives not only serve as a means to assert agency for persons whose control has been diminished but also provide 'wounded' storytellers with a means to reshape their identities" (Sharf & Vanderford, 2003, p. 21).

Narrative and Deception in the Context of North Korea

The use of narratives extends across all contexts of communication—from government-produced stories about national heroes, to memoir writing, and to the stories we tell family and friends, for example. Likewise, as we've learned throughout much of the current handbook, deception also cuts

across virtually all contexts of human interaction. North Korea in particular has built "a strong track record of seeking to deceive both its people and the wider international community" (Murphy, 2014, p. 49) and has used narrative as one of the principal vehicles to do so (Byman & Lind, 2010). The use of deception in North Korea has emerged through various forms, including propaganda via state-controlled media to promote nationalism, cover-ups of human rights violations, and scapegoating of external opposition. Governmental and non-governmental organizations, journalists, and scholars observing North Korea from the outside have pointed to deception primarily in public and mass contexts. Further, there is a dearth of understanding about how deception emerges in the daily lives of individuals living in North Korea, as well as how communication is used by individuals who decide to escape the country and resettle elsewhere. For the purpose of clarity, we define deception as any message that conceals or misrepresents the truth, and contend that it is a powerful force often used to control individuals' thoughts, emotions, and behaviors.

Investigating the nature of deception in various contexts is important for a multitude of reasons. In particular, "Exploring the manifestation of deception in different contexts provides an overall understanding of the concept in its myriad forms, as a deliberate tool directed toward the accomplishment of a specific outcome, an unintentional product of a situation, or an object of analytic inquiry" (Carter, 2014, p. 265). Further, studying deception in contexts that involve humanitarian issues has the potential to give voice to those who have been oppressed and muted, and may work toward further illuminating the nature and effects of human rights abuses. Thus, we endeavor in the current chapter to examine narrative and deception in the context of North Korea, but from the specific perspective of North Korean refugees.

North Korean Refugee Memoirs

A memoir, according to Bailey and Hancock (2014), "is a category of autobiography that focuses on an author's experience rather than his or her entire life" (p. 654). Memoirs of any subgenre (e.g., athletic, travel, addiction) are of course for public consumption and therefore rely on having an interesting, engaging story to tell. Naturally, then, the question of the accuracy in memoir writing is a valid concern. Regarding the subgenre of memoirs written by North Korean refugees, Cussen (2016) outlines various arguments made about their validity and ultimately attests to their value "for what these memoirs tell us of North Korean ground realities, of the horrors of sex and bride trafficking in China's northeast, of the psychological challenges faced by North Koreans who succeed in escaping into the free world—as well as for what they tell us of human pluck, determination, and resilience" (p. 150). In a similar vein, Lankov's (2015) point on North Korean refugee testimony and memoirs is summed up in the article's subheading: "Past treatment of

Soviet, Cambodian atrocity survivors highlights dangers of disregarding defectors' stories." Haggard and Noland (2011) illuminate that "North Korea is a notoriously closed society that not only seeks to control the flow of information into the country, but exercises tight control over information flowing out as well. It is nearly impossible to conduct direct research on any aspect of North Korea" (pp. 3–4). They provide a concise conclusion on the matter—although there are natural disadvantages to self-reported data, the use of "Memoirs…and interviews with refugees provide an important window into life in North Korea" (p. 4).

Lastly, memoirs are part of popular culture, which in and of itself has merit in analyzing and understanding. In her discussion of globalization and social justice, Sorrells (2012) remarks, "Popular culture can function as a platform for discussion or as an initiating force for social change" (p. 156). In the case of the current study, these memoirs serve as an initial platform for understanding how refugees manage their lives in North Korea and navigate their escapes using various forms of deception and truth telling. Doing so helps us better understand the lived experiences, or truths, of some of the world's most repressed populations and is an invitation for others to work for social change. North Korean refugee memoirist Jin-sung Jang (2014) remarks:

> North Korean exiles are a living testament that there does exist a difference between freedom and tyranny. Their stories are not merely a vehicle to evoke pity. They cry for justice on behalf of all those who have died without a voice and who have been buried alive with the world as their dumb witness. Their insistent voices are the triumph of humanity, having survived a brutal struggle with a despot. (p. 313)

Memoirs written by North Korean refugees commonly attest to life in North Korea, escape, and re-settlement. Docan-Morgan (2018a) remarks that there is a "growing genre of memoirs written by North Korean refugees" (p. 120). Beyond traditional paperback and hardcover availability, their popularity is seen via sales as unabridged audio downloads (e.g., via iTunes) and eBooks (e.g., via Kindle). Some memoirs, including Jin-sung Jang's (2014) *Dear Leader: My Escape from North Korea*, have received notable attention as "international best sellers." Others, for example, Hyeonseo Lee's (2015) *The Girl with Seven Names: Escape from North Korea*, made the *New York Times* Best Seller List, which many consider as the preeminent list of best-selling books in the US. Several have been translated into multiple languages, further reaching international audiences—soon after Yeonmi Park's (2015) publication of *In Order to Live: A North Korean Girl's Journey to Freedom*, it was "translated into 15 different languages and was released in 18 different countries" (Park, 2016). North Korean defector memoirs have gained such popularity no doubt because of the extraordinary and shocking—yet true—nature

of the stories they tell, opening a window to the everyday lives so vastly different from what the majority of readers might have known or imagined could exist anywhere in today's world.

Shirly Lee, a notable author, as well as translator for Jin-sung Jang's (2014) memoir, illuminates in the translator's note at the end of Jang's memoir an all too forgotten point about outsiders' collective knowledge of North Korea, reminding us that "in order to make sense of North Korea's present, you have to know its past" (p. 320). She goes on to say,

> Particularly, you have to recognize its persistent dualities—between words and deeds, propaganda and reality, and the manner in which these dualities work for the outsider versus the insider. Without appreciating this, North Korea will remain inscrutable and our exchanges cyclical. (p. 320)
>
> The North Korean regime's doublespeak and opacity are two of its crucial pillars of power. Regardless of whether the world could not see through those façades, or was reluctant to do so, Mr. Jang's memoir reveals that understanding North Korea's past and its persistent dualities is both the key to clarifying its present and to unlocking changes to come. (p. 321)

A further consideration to bear in mind is that many North Korean refugee memoirists address issues of accuracy in their accounts. Early in her memoir, Eunsun Kim (2015), states: "Everything recounted in this book is true. However, to protect the members of my family who will remain in North Korea, I am writing under a pseudonym, and other names and details have changed" (no page number indicated). Other memoirists make similar statements early on in their memoirs. Sungju Lee (2016) offers his readers important context: "*Every Falling Star* is my childhood story based on my memories of events as they occurred at that time. Please note that these were my childhood memories when I was a street boy, suffering from trauma, malnutrition, and starvation as well as sleep deprivation…This is my story as I remember it" (p. 309). Kim Suk-Young, who served as transcriber and translator for Yong Kim's (2009) memoir, states in the preface, "The result is a narrative told in a straightforward but honest voice, interwoven with infrequent emotional reflections, that recounts the events as they happened to a man who suffered the unimaginable." In providing context for her memoir, Yeonmi Park (2015) makes a useful statement that pertains to the majority of memoirs analyzed for the current study:

> The country I grew up in was not like the one my parents had known as children in the 1960s and 1970s. When they were young, the state took care of everyone's basic needs: clothes, medical care, food. After the Cold War ended, the Communist countries that had been propping up the North Korean regime all but abandoned it, and our state-controlled economy collapsed. North Koreans were suddenly on their own. (p. 15)

These reflections point to the admission that the traumatic nature of events experienced by a large number of North Koreans, including many of the memoirists, as well as the passing of time and a need to protect individuals remaining in the country can result in discrepancies. However, the sheer volume of testimonies about life in North Korea which has now come to light affirms patterns of daily life, as well as patterns of abuse that are now so well known as to be irrefutable.

The use of self-reports from North Korean refugees is common in academic research. Refugees are frequently the primary source of data in studies across fields including anthropology, sociology, political science, and social psychology, among others (see Chung, 2008; Jeong & Kim, 2016; Ko, Chung, & Oh, 2004; Song & Bell, 2018). Further, the United Nations Human Rights Council, which established the Commission of Inquiry on Human Rights in the DPRK, discussed in our introduction, relied largely on self-reported refugee accounts (their full report, including methods, findings, and recommendations, is available online; see Commission of Inquiry, 2014). Human rights documentation groups in South Korea, as well as the newly established Human Rights Record Center within the South Korean Ministry of Unification and the Seoul United Nations Office of the High Commissioner for Human Rights work with North Korean defectors to gather testimonies covering a wide range of information relevant to the regime's functioning and the human rights abuses suffered by the people. In the absence of direct access to the country without extreme restrictions on movement, the knowledge of North Korean refugees is therefore our best resource in uncovering life in the country, including information the regime would rather keep hidden.

SYNTHESIS AND STATEMENT OF PURPOSE

This chapter has thus far sought to describe the recent historical context that has resulted in the North Korean refugee phenomenon, highlighting the unique and grave circumstances that surround the creation of the memoirs to be examined. The work of the memoirists, especially if published in English, continues to achieve international popularity, while those who do not publish memoirs nevertheless provide vital information for research and media coverage on the inner workings of the state and society, signifying the importance of defector testimony in the overall narrative on North Korea. The current study provides a rich understanding of North Korean refugee identity and gives voice to individuals who have experienced the oppressive conditions of North Korea, their often dangerous escapes via multiple countries, and nonlinear or frequently difficult times of adjusting to a new society. Stated perhaps most astutely by a North Korean refugee, "if you want to survive in a society like North Korea, you have to be able to deceive yourself and

others" (in Kirkpatrick, 2014, p. 26). The purpose of the analysis below is to understand how North Korean refugees discuss the topics of truth and deception in their published memoirs. The intention of this exploration is largely to spread awareness about and give voice to individuals who have experienced the oppressive conditions of North Korea, their perilous escapes, and often challenging times of readjustment.

METHOD

The sample for the current study consisted of memoirs written by North Korean refugees. Inclusion criteria were that the texts had to be firsthand written refugee accounts (i.e., refugee authored), available for public consumption, and published in English between the years 1990 and 2018. The search for published memoirs was conducted using various search engines (e.g., academic and non-academic databases, Amazon.com) and relevant keywords (e.g., "memoir," "North Korea"). All of the memoirists were born in North Korea and defected or escaped for reasons including hunger, attempting to find family members who had left the country previously, fleeing because of accusations of political wrongdoing, and/or eminent punishment. The search resulted in 11 published memoirs. See Table 51.1 for a complete list.

All language pertaining to or referring to *truth* (e.g., "tell the truth," "the weapon I wield is truth"), *lying* ("the first time I lied," "I won't lie"), and *deception* (e.g., "Everything I learned was a lie to deceive the people," "North Korea uses dialogue as a tool of deception") was recorded verbatim from each memoir, noting the book title, page number, and relevant contextual details pertaining to each quotation. After data collection was complete, we utilized thematic analysis (TA), "a method for systematically identifying, organizing, and offering insight into patterns of meaning (themes) across a data set. Through focusing on meaning across a data set, TA allows the researcher to see and make sense of collective or shared meanings and experiences…This method, then, is a way of identifying what is common to the way a topic is talked or written about and of making sense of those commonalities" (Braun & Clarke, 2012, p. 57). We employed TA's key phases, including reading and rereading the textual data in an immersive, exploratory process to become closely familiar with the authors' accounts of their lives. After, we worked independently to generate initial codes and search for themes, looking for meaningful patterns. We then worked collaboratively, identifying areas of similarity and overlap, comparing initial codes and themes. This process involved constant comparative analysis (Strauss & Corbin, 1998), as we compared examples or extracts for similarities and differences while defining, refining, and naming themes. Throughout our coding, analysis, and write-up, we utilized Braun and Clarke's (2006) "15-point checklist of criteria for good thematic analysis" (p. 96).

Table 51.1 List of North Korean refugee memoirs analyzed in the current study

Author(s)	Memoir title	Copyright/publication date	Pages	Publisher
Sungju Lee, with Susan Elizabeth McClelland	Every Falling Star: The True Story of How I Survived and Escaped North Korea	2016	314	Amulet Books
Yeonmi Park, with Maryanne Vollers	In Order to Live: A North Korean Girl's Journey to Freedom	2015	273	Penguin Books
Joseph Kim, with Stephan Talty	Under the Same Sky: From Starvation in North Korea to Salvation in America	2015	274	Houghton Mifflin Harcourt
Eunsun Kim, with Sébastien Falletti	A Thousand Miles to Freedom: My Escape from North Korea	2015	228	St. Martin's Griffin
Hyeonseo Lee, with David John	The Girl with Seven Names: A North Korean Defector's Story	2015	304	William Collins
Jin-sung Jang	Dear Leader: My Escape from North Korea	2014	339	37INK/Atria
Lucia Jang, with Susan McClelland	Stars Between the Sun and Moon: One Woman's Life in North Korea and Escape to Freedom	2014	280	W. W. Norton & Company
Yong Kim, with Suk-Young Kim	Long Road Home: Testimony of a North Korean Camp Survivor	2009	168	Columbia University Press
Hyok Kang, with Philippe Grangereau	This is Paradise! My North Korean Childhood	2007	204	Abacus
Chol-hwan Kang, with Pierre Rigoulot	The Aquariums of Pyongyang: Ten Years in the North Korean Gulag	2001	238	Basic Books
Soon Ok Lee	Eyes of the Tailless Animals: Prison Memoirs of a North Korean Woman	1999	160	Living Sacrifice Book Company

Findings and Discussion

Our analysis of how truth and deception are discussed in North Korean memoirs yielded 3 overarching themes: (1) discovering and dealing with propaganda, including two subthemes of *worshiping leaders as gods* and *questioning government narratives*; (2) deceiving as a means of survival, including four subthemes of *feigning adherence to propaganda*, *balancing suspicion and secrecy in public contexts*, *balancing suspicion and secrecy in private contexts*, and *deceiving the self*; and (3) finding the truth and living to tell it, comprising the three subthemes of *discovering new realities via foreign media*, *fighting to tell the truth*, and *gaining voice*, *exposing truth*, *and advocating for human rights*. Below, we discuss these themes in more detail, include representative examples from the texts, and at times, and within little space, draw links to relevant information and literature.

Discovering and Dealing with Propaganda

The first primary theme focuses on memoirists' discussion of and dealing with government propaganda, a nearly inescapable reality of North Korean society. Propaganda is defined as "the deliberate, systematic attempt to shape perceptions, manipulate cognitions, and direct behavior to achieve a response that furthers the desired intent of the propagandist" (Jowett & O'Donnell, 2012, p. 7). Extending this definition of propaganda, Patterson, Milburn, and Monteiro (2014) remark, "deception and lying are used as the means of 'shaping,' 'manipulating,' and 'directing'" (p. 418). We found two subthemes for *discovering and dealing with propaganda*, including *worshipping leaders as gods* and *questioning government narratives*.

Worshipping Leaders as Gods

A common theme memoirists recalled during childhood was that their leaders—the generational lineage of the Kim dynasty—were and continue to be deified as gods. Chol-hwan Kang (2001) remarks about his childhood, "I had been made to believe—and had indeed wanted to believe—that the Democratic People's Republic of Korea was the best country in the world. I looked up to Kim Il-Sung as a god" (p. 67). Similarly, reflecting on his learning of Kim Il-Sung, Sungju Lee (2016) shares, "I wanted to be brave and magical, just like him. He was everyone's idol…He was part God too" (pp. 6–7). Hyok Kang (2007) states in his memoir, "To convince us that Kim Junior was a formidable idealist, we were told that as a child he climbed trees in order to catch rainbows, which he finally succeeded in doing…A kind of demigod then! …And I can assure you that at the time, I swallowed it whole" (p. 58). Memoirists also discussed various rituals: "Like all North Korean families, mine kept a shrine on our wall to the Great Leader and his wife. The first thing my father did in the morning…was to take a cloth and carefully dust both of their portraits…You could be sent to a prison camp for

allowing dirt to gather on Kim Il Sung's portrait, or for putting it behind cracked glass" (J. Kim, 2015, p. 11). Hyeonseo Lee (2015), a child when Kim Il-sung died on July 8, 1994, mentions:

> the Great Leader, the father of our nation, was dead…Incredible as it may sound now, it had never occurred to me, or to many North Koreans, that this god-king, so powerful that he could control the weather, might die. He was flawless and almighty. He existed so far above humankind that a part of me didn't want to think he was real. We did not even think he needed to sleep or urinate. (p. 71)

Of course, it was not only children who believed in the demigod nature of the Kim dynasty. In his memoir, for example, Yong Kim (2009) mentions, "I felt like an idiot for having given my life for the Great Leader everyone was brain washed to believe was a living god" (p. 76).

Questioning Government Narratives

Memoirists also spoke of the stories they were told during their youth as having dubious veracity or being outright deceptive. Joseph Kim (2015) elaborates on the stories he learned and their truthfulness: "Like every family we knew, we had a copy of Kim Il Sung's memoirs. They came in an eight-volume set…There were many action scenes and battles, and illustrations of Kim Il Sung and his brethren bayoneting the enemy…Whether it's true or not is another story" (p. 76). This questioning of truth was widely evident, especially in memoirists' youth and teenage years. Sungju Lee (2016) mentions, "History—or what I now call propaganda—was often the first, fourth, and final subject of the day, and the lessons almost always began with the same introduction" (p. 6). Later, Lee tells of a conversation, where his close friend Myeongchul mentions, "'Folklore has a funny way of becoming truth. If we didn't have folk stories, we might start to question our lives, our governments, our world…We might start…thinking for ourselves'" (p. 176). Later in Lee's memoir, at the point where he was forced to live on the streets and fend for himself, he recounts the words of a gang leader whom Lee befriended:

> My gang and I don't believe in Joseon [North Korea], because it lies to us. It says Joseon is a paradise and children its kings and queens. But children are dying from terrible starvation and diseases. Kings and queens don't die like this. The military are thieves…They don't protect people; they steal. I don't believe in the army, not anymore. (p. 204)

Jin-sung Jang (2014), who grew up in an elite family, received private music education, and enjoyed playing the piano, shares his questioning of government narratives as a teenager:

…our [school] music teacher punished me for my [pianistic] deviation by humiliating me in front of the class, making an example of me as someone who knew nothing whatsoever about music. In my heart, though, I believed it was the school—not me—that lacked an understanding of music…As a result, I could not stop myself from beginning to doubt everything else the school taught us to regard as the most accurate and objective form of knowledge, whether this took the form of the revolutionary history of Kim-Il Sung, linguistics, or any other subject. (p. 31)

Deceiving as a Means of Survival

The second primary theme we address concerns how memoirists discuss deception as a means of survival for living in North Korea, including *feigning adherence to propaganda*, *balancing suspicion and secrecy in public*, *balancing suspicion and secrecy in private*, and *deceiving the self*. Many of these strategies also fall within the realms of life-saving lies (e.g., Akhtar, 2009) and high-stakes deception.

Feigning Adherence to Propaganda

Memoirists commonly wrote of instances of feigning adherence or faking devotion to the North Korean government and particularly to its leadership. Hyeonseo Lee (2015) writes of pretending to be visibly saddened and distressed when her classmates gathered in front of the school building the morning after Kim Il-Sung's death: "If I didn't cry like everyone else, I'd be in trouble. So I rubbed my face in false distress, surreptitiously spat on my fingertips, and dabbed my eyes. I made a gasping noise that I hoped sounded like I was heaving with despair" (p. 72). Feigning adherence was often necessary in the form of public performance. In his memoir, Sungju Lee (2016) reflects: "I realized after spending nearly all my time with Young-bum that when he had chanted for prisoners to be executed, he wasn't doing so because he believed they should be killed. He was putting on a show so the principal and the *so-nyon-dan* manager [a leader of the organization for children that is heavily involved with propaganda] wouldn't think he was a criminal, too" (p. 113).

Balancing Suspicion and Secrecy in Public Contexts

Another common theme memoirists discuss concerns living in a public culture of suspicion and mistrust, which often resulted in careful secrecy and concealment. Hyok Kang (2007) remarks, "In North Korea everyone is suspicious of everyone else, all the time. There are security spies in every work unit, but you never know who they are, or how many" (p. 47). Memoirists commonly discuss their childhood training in writing and orally reciting letters of criticism, illuminating the necessity and ubiquity of suspicion and secrecy. Kang provides a clear illustration:

to succeed, it would have to remain secret. Even our families would have to be kept in the dark, and telling friends was out of the question. (p. 187)

Jin-sung Jang (2014) recounts a similar situation regarding his escape, utilizing secrecy to avoid suspicion from his parents. Jang's case also serves as an example of the complicated, intertwined nature of suspicion, secrecy, and deception, which are also entangled across contexts (e.g., political level and interpersonal level). Working as a high-level psychological warfare and propaganda officer, he made the mistake of removing a South Korean periodical from his work unit, the United Front Department, which is responsible for establishing pro-North Korean groups in South Korea. Making matters worse, Jang lent the periodical to a close friend who subsequently lost it in short time. The periodical

included a biography of Kim Il-sung and Kim Jong-il written by a South Korean academic who had pieced together their family history, although we were only allowed to know their revolutionary history. It even made mention of the fact that Kim Jong-il had mistresses…it was considered treason of the most serious degree to have shared this information. (p. 70)

Because of the importance of the periodical, Jang and his friend knew authorities would attempt to find and confront them quickly. Thus, they had realized they had to escape the country or would be caught and severely punished, if not executed. The entangled matter of secrecy, suspicion, and having to deceive to survive—even in family contexts—is further illustrated in Jang's departure from his family:

I felt sick that my mother and father must live out their remaining days in a world from which their only son disappeared. Yet I could not say good-bye to them. They would not let me go if they learned my plan. They would kill themselves first. Once I left the country and officially became a missing person, I knew how the Ministry of State Security would interrogate them. If they so much as suspected that my parents had been aware of my intention to escape, they would be convicted of assisting a traitor. It was far better for them to face the authorities in complete innocence. As I walked into the living room, they both questioned me at once. "What are the sunglasses for?" "My eyes are a bit sore." I managed to make up an excuse. …My father intervened. "Let him alone, he'll be late for work."…I quickly crossed the living room and made my way towards the front door. Only when I reached the threshold did I steal a look behind me. I longed to see my parents one more time…As soon as I left the house, my tears erupted in bitter sobs, I knew the Worker's Party could take away my right to life, but it had also taken away my right to say good-bye to my family, and I had to deceive them to the end. (pp. 73–75)

Deceiving the Self
The final subtheme of *deceiving as a means of survival* concerns self-deception. One popular definition states in part that self-deception "is a motivated

unawareness of conflicting knowledge in which threatening knowledge is selectively filtered out from consciousness as a psychological defense…" (Starek & Keating, 1991, p. 146). Early in her memoir, Yeonmi Park (2015) states, "I believed that, somehow, if I refused to acknowledge the unspeakable past, it would disappear. I convinced myself that a lot of it never happened; I taught myself to forget the rest" (p. 5). Later, she states bluntly, "…we North Koreans can be experts at lying, even to ourselves" (p. 54). She provides additional context: "North Koreans have two stories running through their heads at all times, like trains on parallel tracks. One is what you are taught to believe; the other is what you see with your own eyes" (p. 53). She goes on to provide various examples, such as "It is how you can recite the motto 'Children are King' in school, then walk past the orphanage where children with bloated bellies stare at you with hungry eyes," and "The frozen babies that starving mothers abandoned in the alleys did not fit into my worldview, so I couldn't process what I saw. It was normal to see bodies in the trash heaps, bodies floating in the river, normal to just walk by and do nothing when a stranger cried for help" (p. 54). She elaborates on the nature of actively filtering out information and its effects:

> There were so many desperate people on the streets crying for help that you had to shut off your heart or the pain would be too much. After a while you can't care anymore. And that is what hell is like. (p. 55)

While Yeonmi Park's (2015) example is useful for understanding why one would engage in self-deception due to societal distress, memoirists also wrote repeatedly of deceiving the self in interpersonal and family contexts. Lucia Jang (2014), born in the 1970s, provides a useful example of self-deception. As a young woman, Jang worked in a factory. There, she met Myungin, a man she thought was courting her, but instead raped her. She was forced to marry him when she found herself pregnant and continued to endure his abuse, using self-deception as a psychological defense mechanism:

> Myungin could do anything he wanted. As he dealt the blows, I knew I was nothing to him except a pitiful girl whose bride money was needed for his family's debts. He had never adored me. He had never cared. What hurt me most was that I had known this truth from the beginning. I just didn't want to see it. (p. 118)

Finding the Truth and Living to Tell It

The final primary theme concerns finding new truths and exposing them to the outside world. Three subthemes include *discovering new realities via foreign media*, *fighting to tell the truth*, and *gaining voice, exposing truth, and advocating for human rights*.

Discovering New Realities via Foreign Media

A common discussion point among memoirists was not only that they consumed foreign, illegal media, but their reflections on how doing so changed their thinking about the reality in and beyond North Korea. Hyok Kang (2007) remarks:

> As we were quite close to the Chinese border, we were able to pick up the Beijing channels. That was totally and utterly forbidden but we did it anyway, at night, with the curtains drawn. Chinese television gave us an absolutely incredible view of the world. There were cars everywhere, rich people who ate all the time and delicious-looking food, buildings that looked like mirrors, lovely homes piled high with household appliances and electric gadgets…Chinese television looked a hundred times truer than our one channel… (p. 43)

In a similar vein regarding media and reality, in Yeonmi Park's (2015) discussion of the contradictory narratives North Koreans have running though their heads, she states:

> It is how you can believe that North Korea is a socialist paradise, the best country in the world with the happiest people who have nothing to envy, while devouring movies and TV programs in enemy nations enjoying a level of prosperity that you couldn't imagine in your dreams. (p. 54)

Chol-hwan Kang (2001) provides another example: "Listening to South Korean radio had to be done with extreme caution" (p. 185). He recalls listening to Christian programs on the Korean Broadcasting System, as well as the Voice of America, and reflects, "Listening to the radio gave us the words we needed to express our dissatisfaction. Every program, each new discovery, helped us tear a little freer from the enveloping web of deception" (p. 186). Kang's example aligns with Fahy's (2015) notable work with refugees, which in part describes the process of North Koreans' social and psychological strategies for coping with the regime through their personal experiences of suffering, enlightenment, and disappointment.

Although foreign media is illegal and can come with heavy consequences if one is found possessing, consuming, or distributing it (Korea Institute for National Unification, 2016), many analysts and North Korean refugee memoirists who discuss the topic tend to agree with Chol-hwan Kang's (2001) perspective about foreign media having a clear role enabling North Koreans to "tear freer." Numerous authors and advocates argue that the continued inflow of foreign media into North Korea is the primary way to affect the social consciousness of North Koreans, and that foreign media has already generated irreversible changes in the country. Baek (2016), who provides one of the most recent and comprehensive analyses of the topic (Docan-Morgan, 2018b), discusses the ways in which forbidden information is spread through

gossip, freedom balloons, radio, and USBs. She admits that although foreign media does not work as a magic bullet to liberate individuals, it "may be instrumental in someday bringing down one of the most brutal and repressive regimes in modern history" (p. x). She contends, "more information will drive more social and cultural changes" (p. 196).

Fighting to Tell the Truth

Many memoirists wrote about their grueling experiences to stay alive in order to ultimately inform others about their realities in North Korea, as well as their escapes in search of freedom. This subtheme appears in several ways. Some memoirists spoke of their relentless will to stay alive *inside* of North Korea with the sole purpose of escaping and sharing their realties with the rest of the world. Other memoirists spoke of realizing the regime's lies *during their escape and quest for freedom*, indicating a desire to tell their truths to the world. And numerous memoirists also spoke of an additional fight to tell the truth—the unexpected fight to be believed by others once they found freedom.

Soon Ok Lee (1999), for example, is one memoirist who writes about her drive to remain alive in North Korea, even under brutal conditions, with the singular purpose of escaping and sharing her realties with the rest of the world. She describes her life, working as a well-respected supervisor at a material distribution center; however, because she refused to satisfy the greed of a government officer, she ended up enduring six years of inhumane treatment in prison, starting in 1986. Her experiences align with aforementioned themes (e.g., *discovering and dealing with propaganda*): "Sometimes people see the truth. That happened to me. I saw some of the absurdity of the North Korean government before I was sent to prison, but my childhood training in Kim Il Sung's doctrine kept me from seeing the truth. It wasn't until I suffered from the injustice that I began to change my mind. Then my eyes were opened to the system I was living in" (p. 121). In her memoir, she wrote recurrently about her will to survive in order to inform the outside world. For example: "I am disgusted by the lies of the North Korean government. I once truly believed that North Korea was the paradise of the universe, but it is really the den of evil. Everyone in the world will see the reality of hell when the Korean government is torn down" (p. 142). Related to the subtheme of *fighting to tell the truth*, Lee (1999) and multiple memoirists describe a common response to their stories once they find freedom—disbelief in their experiences. In her conclusion, Lee states:

> When I first began to testify of the brutality in North Korea, no one believed me. Someone told me, "No way! How could people survive in such an environment!" Perhaps it is natural that people who have not suffered like this think I'm exaggerating, but I am sad to admit that this is true and is happening right now. (p. 154)

Lee's comments point to the reality that some refugees fight to escape to freedom in order to inform the world about life in North Korea, but also experience a continued fight once in freedom—the *fight to tell the truth* and *be believed*. Hyok Kang (2007), who escaped from North Korea when he was 12 years old, spent four years in China, and arrived in South Korea at the age of 16, reflects on his childhood in the South: "When I tell children of my own age in South Korea what life is like in North Korea, most of the time they don't believe me" (p. xi). Like Soon Ok Lee (1999) and Hyok Kang (2007), Chol-hwan Kang (2001) tells of his doubters, including media:

> A month after our arrival, we were brought to the Seoul Press Center to be interviewed by several dozen journalists…I had been through so many awful things, and these people, who had lived their whole lives swaddled in perfect comfort, were looking skeptically down their noses at me!…I found the journalist from the newspaper *Hangyore* particularly irritating. What place did his skepticism leave for the victims? Millions of people were dying or suffering from hunger, an entire population was being deprived of its freedom, and his only concern was our credibility…I decided to speak. "If you don't want to believe us, go to the North! Do you think we risked our lives so we could come here and lie?" (pp. 223–224)

Yeonmi Park (2015) also writes about dealing with skepticism from the media. Once Park (2015) began sharing her story, often in English, she realized that some people were "keeping score" of every detail she shared (p. 263). In her memoir, she writes of having to make crucial decisions about what to share and what to keep private, especially as her story gained intense media interest. Park shares that ultimately, her mother "wanted people to know why we had to escape, and what happened to North Korean women who were sold in China…'If you don't speak up for them, Yeonmi-ya, who will?' she said. My sister agreed" (p. 264). Park mentions, "I would soon discover that to be completely free, I had to confront the truth of my past" (p. 257). She tells of choosing to disclose more of her experiences, including being trafficked with her mother and seeing her mother being raped by a Chinese broker. As Park's story gained more attention, the North Korean government began watching her closely. She remarks:

> In early 2015, the regime uploaded two separate videos calling me a liar and a "human rights propaganda puppet." They had sifted through my interviews and attacked me for supposed inconsistencies in my quotes. When the regime couldn't dispute what I said, they invented lies about me and my family…Worst of all, they paraded my relatives and former friends to denounce me and my family. (pp. 264–265)

Fighting to tell the truth appears to have no boundaries. Eunsun Kim (2015), who grew up in North Korea during the famine remarks: "we in North Korea

did not have any information about the rest of the world, other than what was fed to us through the state's propaganda, which always emphasized that it was far better to live here than in the chaos of the capitalist world. We grew up in one big lie, but I didn't know it then" (p. 50). She elaborates, indicating a desire to share the realities of North Korea: "It was only later, at the end of our perilous journey in search for freedom, that my eyes were opened to the subservience of our lives in North Korea and that I began to understand the horror of that inhumane regime. Today, I can only denounce the regime's crimes, because I am safely in South Korea. And here, at long last, my stomach is full" (p. 51). Now in South Korea, Kim speaks about her experiences living in North Korea, escaping to freedom, and her adjustment (Worrall, 2015).

Similar to Eunsun Kim (2015), Jin-sung Jang (2014) also experienced a turning point during his journey for freedom. After fleeing North Korea, Jang was struggling to survive in China. While there, his close friend with whom he escaped North Korea, Young-min, had died. Jang remarks:

> From that moment on, I was no longer a fugitive. I was no longer fleeing out of terror, but fighting for my freedom, so that I could expose the lies of Kim Jong-il. I wasn't afraid to die if I died a free man, and this released me from fear. (p. 293)

Later, Jang remarks:

> The North Korean regime has not finished with its persecution of me. It not only makes secret attempts to find and harm me physically, it also threatens me openly through its media. In June 2013, for example, the Ministry of People's Security published an official statement through the North Korean state news organ, KCNA, saying it would "remove my existence from this universe." The tyranny of Kim has now been inherited by a third generation. This is why my peace lies in waging war against despotism, until our people are freed. Without that, my privilege of freedom would be no more than selfishness. (p. 314)

Gaining Voice, Exposing Truth, and Advocating for Human Rights

All texts examined in the current study purposefully expose memoirists' truths or realities about life in and escaping North Korea. They also directly advocate for human rights for North Koreans. Indeed, before and/or after their memoirs were published, the majority of these memoirists are human rights advocates—giving speeches, testimonies, and media interviews. Some have created their own organizations helping refugees or have expressed plans to do so in the future. Within their memoirs—often in the preface, epilogue, or afterword sections—memoirists discuss their activities at the time of writing. At the end of Jin-sung Jang's (2014) memoir, for example, he discusses his journey exposing the truth. He was chosen to represent North Korea in

exile, among several other poets in exile, for the World Poetry Summit at Poetry Parnassus, as part of the Summer Olympics held in London in 2012. He reflects that the event "strengthened my resolve to declare the truth about North Korea through the written word" (p. 311). Notably, Jang also created *New Focus* (www.newfocus.co.kr). In Jang's words, it is

> the first news organization run by North Korean exiles. I named it thus for two main reasons: in the hope that North Korea could pursue a new vision; and to show the outside world that there is a way of understanding North Korea beyond the way that existing frameworks of interpretation or government agendas allowed it. I wanted the knowledge and experience of North Koreans to be taken seriously into account. As I've written elsewhere: "[Back home] there are two North Koreas: one real and the other fiction. After my defection, I recognized the existence of a third North Korea: a theoretical one, one constructed by the outside world…" (p. 311) …our guiding principle from the start has been: "Don't worry about going faster than those who have had a head start; worry only about being more honest. It may take a long time for the truth to come to light, but it will remain long after the lies have faded." (p. 316)

Chol-hwan Kang (2001), who spent 10 years in North Korea's Yodok concentration camp, provides an updated preface for the revised edition of his memoir in 2005. In South Korea, he worked as a staff writer specializing in North Korean affairs for the *Chosun Ilbo* newspaper. He reports having "met and reported on approximately 500 North Korean refugees and defectors, those on the run in China and those who found freedom in South Korea" (p. vii). In 2007, he founded the North Korea Strategy Center (NKSC), which is active in South Korea and the US. According their website, "As a defector-led organization, NKSC US believes that North Koreans are leading change in North Korea. We accelerate this people powered change by providing a platform for North Korean voices. Our programs empower North Koreans within the country with access to information, while supporting defectors outside of North Korea with leadership development programs and international support networks" (North Korea Strategy Center, n.d.).

Memoirists have also utilized their voices, exposed truth, and advocated for human rights by partnering with existing organizations. For example, Hyeonseo Lee (2015) and Joseph Kim (2015) write about their experiences spreading awareness by giving TED Talks. Yeonmi Park (2015), who perhaps has received the most international media attention of the 11 memoirists, writes about giving a multitude of speeches and interviews about her experiences. Both Yeonmi Park (2015) and Joseph Kim (2015) praise Liberty in North Korea (LiNK). Park, for example, addresses the LiNK staff: "you taught me what it means to be a spokesperson for the North Korean people… and become a better advocate for freedom" (p. 270).

Certainly, the decision to publish a memoir exposes one's own truths and, within the current context, advocates for North Korean human rights.

Eunsun Kim (2015) addresses this clearly, as well as her partnership with an existing organization:

> I undertook the writing of this book with a mission that I hold dear to my heart: providing witness testimony to the situation in North Korea, and helping to alleviate the burdens of my people, who are oppressed by a totalitarian dictatorship. That's why for now, I work for an NGO based in Seoul. The Citizens' Alliance for North Korean Human Rights (NKHR) tries to mobilize world leaders to change the fate of the Korean peninsula north of the 38th parallel, and to help North Korean defectors who have taken refuge in Seoul. (p. 226)

Refugees have been instrumental to the important work of the Citizens' Alliance for North Korean Human Rights, which advocates internationally to raise awareness of the North Korean human rights problem and was a major contributor to the successful campaign to establish the UN COI. Related, Sungju Lee (2016) writes, "In the spring of 2015, I became the consultant for the rescuing team of Citizens' Alliance for North Korean Human Rights, a nonprofit group the helps defectors trapped in China. I speak around the world, raising awareness and money to rescue North Koreans in China" (p. 307). Similarly, in the preface of Hyok Kang's (2007) memoir, Philippe Grangereau writes of meeting Kang at an event organized by human rights organizations: Kang "was invited by the People in Need Foundation (PINF), which was holding the fourth conference of the South Korean NGO 'North Korean Human Rights' (NKHR)" (p. xi).

North Korean refugees who have written memoirs, as well as many who have not, have partnered with other existing organizations in varying capacities. Some participate in mentoring programs to support young defectors newly arrived in South Korea, work as reporters for *Daily NK* (a North Korea-focused news site owned by Unification Media Group), or work as researchers and contributors to projects aimed at documenting ongoing human rights abuses in North Korea.

Refugees have also worked with the NGO Now Action & Unity for Human Rights (NAUH). Seong-ho Ji, a North Korean defector and President of NAUH, states the following on their website's welcome message: "We have spearheaded the effort for alerting the public to the reality of North Korea's human rights conditions, organizing campaigns calling for unification, hosting cultural exchanges between South and North Korean young adults, participating in radio broadcasts that relay news of freedom for North Korea, and helping rescue operations of North Korean refugees" (NAUH, n.d.). Further, the Database Center for North Korean Human Rights (NKDB) and the Transitional Justice Working Group (TJWG) both work with North Korean refugees to document human rights abuses for use in future investigations into crimes committed in North Korea, as well as to map patterns of abuses.

A Call for Purposeful Research and Action

The findings of the current study align with and extend previous work by scholars and practitioners who have found that the Kim dynasty is deified, propaganda is a dominant form of communication in North Korea, and many North Koreans face countless obstacles within their own state and as refugees (e.g., Byman & Lind, 2010; Cha, 2013; Cumings, 2005; Demick, 2010; Haggard & Noland, 2011; Hassig & Oh, 2015; Lankov, 2013, 2017). However, the current analysis takes us a step further, offering an explanation as to how some North Koreans (i.e., memoirists and perhaps an unidentified number of others) are left with the complicated, highly delicate, and potentially dangerous obstacle of questioning government narratives, feigning adherence to the Kim regime, and fighting to tell the truth. Memoirists also brought to light the complicated psychological and communicative acts of balancing suspicion and secrecy in public contexts *and* balancing suspicion and secrecy in private contexts, as well as the coping mechanism of deceiving the self. These highly complex issues undoubtedly take a toll on individuals' well-being and identity—all of which are ripe topics to be explored in more depth for future studies and potential issues to be addressed in relevant social-psychological services offered to refugees.

As scholar-practitioners engaged with issues of human rights, we focus here on practical application, the extension of scholarly knowledge, and improvement of the human condition. Therefore, we end with a call to action aimed at empowering survivor voices, utilizing information to create change in North Korea, and offering future directions for scholars, practitioners, and human rights advocates.

Empowering Survivor Voices

The current study also makes it clear that refugees need more opportunities to gain and express their voices and communicate their realities (if they wish), as they are at a severe disadvantage during and after resettling due to a plethora of documented psychological and cultural challenges. We call upon individuals interested in human rights—whether academics, practitioners, lay persons, clergy, and governmental or non-governmental organizations—to move to action in helping empower survivor voices. This may involve a utilizing a scholar-practitioner model (McClintock, 2004), action research (Chevalier & Buckles, 2013; Lewin, 1946; Reason & Bradbury, 2008), applied research (e.g., Frey & Cissna, 2009), critical grounded theory (Charmaz, 2013), public scholarship (Stoller, 2016), bridging theory and practice via phronetic social science (Flyvbjerg, 2001), community service learning (Bose, Horrigan, Doble, & Shipp, 2014), and/or other practice-based approaches. Direct avenues also include becoming involved with governmental or non-governmental organizations that help refugees gain essential skills, as well as platforms to express their voices, as desired.

Burgeoning research which looks at the issues surrounding transitions from conflict or situations of authoritarian government warns against mistreatment or neglect of survivors of oppressive regimes, drawing attention to key considerations for policy architects, activists, and human rights practitioners who are looking to the future with survivors in mind (Robins, 2011; Robins & Wilson, 2015; McEvoy & McConnachie, 2012). To date, planning in South Korea and internationally for a potential political transition or regime change in North Korea (usually within the framework of inter-Korean unification) has tended to involve North Korean defectors as relatively passive providers of information or witnesses to crimes. Little space is dedicated to refugee agency in the design and planning of institutions and mechanisms that might be created to process what has happened throughout North Korea's difficult history, such as through truth telling, criminal prosecutions, or reparations for victims. Research on rebuilding nations following dramatic transitions finds that institutional approaches led by national elites often marginalize victim constituencies, rather than rooting recovery efforts in an understanding of how mass violations have impacted and transformed affected populations, and what survivors *themselves* perceive will aid most in their recovery from the impact of the violation (Gready & Robins, 2017). Creating opportunities for North Korean defectors to be empowered, if they so desire, will be essential to the success of future peace-building and national restoration, and their active participation can and should be harnessed now. At the time of this writing, research is currently underway on these themes by the Seoul-based Transitional Justice Working Group, in the hope that it will inform current and future engagement with survivors of North Korean human rights abuses as part of the process of dealing with the past.

Information, Communication, and Change in North Korea

The current study also adds to our developing understanding of the importance of information and media that comes into and out of North Korea. In her conclusion on the topic of how the information underground is transforming North Korea, Baek (2016) states, "Civil society organizations and possibly government-agency-powered efforts to increase the flow of information into North Korea may well be the most reasonable, sustainable, cost-effective, and peaceful way of creating positive change inside North Korea" (p. 216). The availability of more information, Baek argues, gives the North Korean people "the agency, self-determination, and knowledge to write their own future and destiny as a nation" (p. 217). Baek invites interested individuals to become involved with organizations that send information into North Korea, and mentions some of the components involved in these processes—"researching best practices from comparative situations, finding and/or creating technologies for dissemination purposes, creating and editing original digital content, fundraising, and more" (p. 225). Although we largely agree with Baek, we would be remiss to not remind readers and practitioners of the

potential consequences for individuals caught with or distributing foreign media in North Korea (see Korea Institute for National Unification, 2016).

Another means of information transmission from the outside into North Korea can occur through international engagement activities such as the delivery of medical and other humanitarian aid, as well as educational initiatives such as the foreign-run and staffed Pyongyang University of Science and Technology, and Choson Exchange, which brings North Koreans to Singapore to study business. Contact with foreign individuals, products, and expertise can have a transformative effect on the view of the outside for the North Koreans who encounter them. The degree to which foreign media and people-to-people contact is transforming North Korea remains to be fully seen, as only time (and additional information) will tell. For practitioners (e.g., NGOs and interlinked networks of actors who push illegal media into North Korea; see Baek, 2016; Docan-Morgan, 2018b), academics, and refugees, we ask: what have been the (verified) effects of people-to-people contact and information flow into North Korea? Has the information flow into North Korea actually created an "underground revolution" as Baek argues, and if so, among whom and how can it safely spread further? Or is it sowing the seeds for change, and what kind of change? In North Korea specifically, (how) can an underground revolution of any type last, and can it lead to more dramatic and wide-reaching change? Will the spread of foreign media in North Korea be a catalyst for an "above ground" revolution, and if so, what exactly is required for this to take place? How, if at all, do history, relevant revolutions, extant research, and theory and practice related to media and diffusion of innovations (see Rogers, 2003) inform us about the best next steps for pursuing reasoned, purposeful, and impactful distribution of information in North Korea? Pragmatically, what types of information and media, technologies, hardware, and diffusion strategies are the most effective and safe?

Future Directions

Experts from a plethora of academic fields can do more to engage with issues faced by North Koreans and refugees. For scholars focused on the areas of deceptive communication, truth, and ethics, we can use our skills and knowledge to confront human rights violations. For example, applied research questions include the following: How can deceptive tactics be used for altruistic purposes (e.g., freeing North Koreans stuck in China)? Regarding a topic nearly absent from the academic literature: What are the more detailed nuances, complications, and successful ways to succeed at life-saving lies? Regarding truth, how can we better understand the experiences of North Korean refugees and help them spread their lived experiences, as they wish? For scholars who are focused on continued theory and research development, there are countless avenues to pursue that will increase our understanding of many of the aforementioned issues—dialectical tensions experienced within personal relationships, identity management, survivor participation, and

decision-making are examples of areas where social scientific, humanistic, and critical approaches will help us better understand the experiences of and challenges endured by refugees.

For scholars dealing with any sort of media, and in any context—marketing, advertising, public relations, propaganda, and politics, for example—what communicative strategies can be created and implemented to better inform the world about human rights violations suffered by North Koreans and the refugee experience? The prominent media narrative on North Korea focuses on nuclear weapons and the political diatribe of state leaders (Kang, 2018), while North Koreans continue to suffer human rights abuses. How can this narrative be amended or refocused on issues of human rights and the realities (i.e., truths) of those who live in North Korea and/or have resettled elsewhere, for example? The potential areas to explore—via research, theory, and practice—are endless and necessary. Our call for purposeful research and action is aptly summed up by Jin-sung Jang (2014), and we believe also applies to North Koreans, refugees, and advocates of human rights: "if the regime has murder, deception and nuclear bombs in its arsenal, the weapon I yield is truth" (p. 314).

Notes

1. There is no single, agreed-upon English spelling or appearance of Korean names. In Korean, the family name commonly appears first, followed by one or two given names. In the current chapter, we hyphenate first names (e.g., Kim Il-Sung) to make it clear which are family and given names. All direct quotations appear as they were originally written; therefore, there are minor differences in how some names appear throughout this chapter (e.g., Kim Il-Sung, Kim Il Sung, Kim Il-sung). However, the majority of North Korean refugee memoirs written in English place given names first and family names second; therefore, we also use this approach.

2. The exact figures for famine deaths remain unclear today. The lowest estimates are at around 450,000, while the highest credible estimates are closer to 1 million, depending on the source (Lankov, 2013, p. 79).

3. For some North Korean defectors, China is a transit country, while others stay indefinitely (Lee, 2017, pp. 34–35). Kim (2012) remarks, "The exact number of North Korean escapees in China is open to debate. The Chinese government's conservative estimate is 10,000; Seoul's calculation is between 10,000 and 30,000; humanitarian organizations put the figure as high as 300,000" (p. 45). Also see Tanaka (2008).

4. The significantly higher percentage of female defectors is due to the greater flexibility women have in leaving the country unnoticed, given that housewives and independent traders (a mainstay of the controlled private markets) will not be missed in the same way as men not turning up for work at government-assigned jobs (Kim, 2013).

5. In this chapter, we use both *defector* and *refugee* synonymously, recognizing the variety of individual preferences expressed by North Koreans themselves in terms of how they choose to identify.

6. In a recent survey of 153 North Koreans defectors living in South Korea on the theme of national identity, the Seoul-based NGO, Transitional Justice Working Group, found that 36.7% of respondents identified themselves as "just Korean," 30% identified as "North Korean," and 25.3% identified as "South Korean" (Transitional Justice Working Group, forthcoming).

7. It is worth mentioning that the publication of a memoir, especially in English with the support of a co-author as many have done, can provide a source of additional income and/or leverage with which to engage in North Korean human rights/anti-regime activism, both of which can be difficult to achieve if there is an absence of other professional skills and experience.

References

Akhtar, S. (2009). Lies, liars, and lying: An introductory overview. In S. Akhtar & H. Parens (Eds.), *Lying, cheating and carrying on* (pp. 1–16). Lanham, MD: Rowman and Littlefield.

Baek, J. (2016). *North Korea's hidden revolution: How the information underground is transforming a closed society.* New Haven, CT: Yale University Press.

Bailey, T., & Hancock, J. T. (2014). Memoirs. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 654–656). Thousand Oaks, CA: Sage.

Bakhtin, M. M. (1981). *The dialogic imagination* (C. Emerson & M. Holquist, Eds., original work published in 1975). Austin, TX: University of Texas Press.

Baxter, L. A. (2011). *Voicing relationships: A dialogic perspective.* Thousand Oaks, CA: Sage.

Bose, M., Horrigan, P., Doble, C., & Shipp, S. C. (Eds.). (2014). *Community matters: Service-learning in engaged design and planning.* New York: Routledge.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77–101.

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper (Ed.), *APA handbook of research methods in psychology, Vol. 2. Research designs* (pp. 57–71). Washington, DC: American Psychological Association.

Byman, D., & Lind, J. (2010). Pyongyang's survival strategy: Tools of authoritarian control in North Korea. *International Security, 35*(1), 44–74.

Carter, E. (2014). Deception in different contexts. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 265–268). Thousand Oaks, CA: Sage.

Cha, V. (2013). *The impossible state: North Korea, past and future.* New York: HarperCollins.

Charmaz, K. (2013). Grounded theory methods in social justice research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (4th ed., pp. 291–336). Thousand Oaks, CA: Sage.

Chevalier, J. M., & Buckles, D. J. (2013). *Participatory action research: Theory and methods for engaged inquiry.* London, UK: Routledge.

Chung, B. H. (2008). Between defector and migrant: Identities and strategies of North Koreans in South Korea. *Korean Studies, 32,* 1–27.

Chung, S., & Seo, J. Y. (2007). A study on posttraumatic stress disorder among North Korean defectors and their social adjustment in South Korea. *Journal of Loss and Trauma, 12*(4), 365–382.

Commission of Inquiry on Human Rights in the Democratic People's Republic of Korea. (2014). *Report of the Commission of Inquiry on Human Rights in the Democratic People's Republic of Korea*. Retrieved from https://www.ohchr.org/EN/HRBodies/HRC/CoIDPRK/Pages/CommissionInquiryonHRinDPRK.aspx.

Committee for Human Rights in North Korea. (2009). *Lives for sale: Personal accounts of women fleeing North Korea to China*. Washington, DC. Retrieved from https://www.hrnk.org/uploads/pdfs/Lives_for_Sale.pdf.

Cumings, B. (2005). *Korea's place in the sun: A modern history*. New York: W.W. Norton.

Cussen, J. (2016). On the call to dismiss North Korean defectors' memoirs and on their dark American alternative. *Korean Studies, 40,* 140–157.

Demick, B. (2010). *Nothing to envy: Ordinary lives in North Korea*. New York: Spiegel & Grau.

Docan-Morgan, T. (2018a). Review of "Every falling star: The true story of how I survived and escaped North Korea" by Sungju Lee. *North Korean Review*, *41*(1), 120–122.

Docan-Morgan, T. (2018b). Review of "North Korea's hidden revolution: How the information underground is transforming a closed society" by Jieun Baek. *Korean Studies, 42,* 94–96.

Fahy, S. (2015). *Marching through suffering: Loss and survival in North Korea*. New York: Columbia University Press.

Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Oxford, UK: Cambridge University Press.

Frey, L. R., & Cissna, K. N. (2009). *Routledge handbook of applied communication research*. New York: Routledge.

Gready, P., & Robins, S. (2017). Rethinking civil society and transitional justice: Lessons from social movements and 'new' civil society. *The International Journal of Human Rights, 21*(7), 956–975.

Haggard, S. H., & Noland, M. (2011). *Witness to transformation: Refugee insights into North Korea*. Washington, DC: Peterson Institute for International Economics.

Hassig, R., & Oh, K. (2015). *The hidden people of North Korea: Everyday life in the Hermit Kingdom* (2nd ed.). Lanham, MD: Rowman and Littlefield.

Human Rights Watch. (2002). *The invisible exodus: North Koreans in the People's Republic of China*. New York. Retrieved from https://www.hrw.org/reports/2002/northkorea/#P64_1170.

International Crisis Group. (2006). *Perilous journeys: The plight of North Koreans in China and beyond*. Retrieved from https://d2071andvip0wj.cloudfront.net/122-perilous-journeys-the-plight-of-north-koreans-in-china-and-beyond_0.pdf.

International Crisis Group. (2011). *Strangers at home: North Koreans in the South*. Seoul/Brussels. Retrieved from https://www.crisisgroup.org/asia/north-east-asia/korean-peninsula/strangers-home-north-koreans-south.

Jang, J. S. (2014). *Dear leader: My ecscape from North Korea*. New York: 37INK/Atria.

Jang, L. (2014). *Stars between the sun and moon: One woman's life in North Korea and escape to freedom*. New York: W.W. Norton.

Jeong, H. O., & Kim, Y. S. (2016). North Korean women defectors in South Korea and their political participation. *International Journal of Intercultural Relations, 55,* 20–31.

Jowett, G. S., & O'Donnell, V. (2012). *Propaganda & persuasion* (5th ed.). Thousand Oaks, CA: Sage.

Kang, C. H. (2001). *The aquariums of Pyongyang: Ten years in the North Korean gulag*. New York: Basic Books.

Kang, D. (2018, July 3). Reasons to be optimistic about North Korea. *East Asia Forum*. Retrieved from http://www.eastasiaforum.org/2018/07/03/reasons-to-be-optimistic-about-north-korea/.

Kang, H. (2007). *This is paradise! My North Korean childhood*. London: Abacus.

Kim, E. (2015). *A thousand miles to freedom: My escape from North Korea*. New York: St. Martin's Griffin.

Kim, J. (2015). *Under the same sky: From starvation in North Korea to salvation in America*. New York: Houghton Mifflin Harcourt.

Kim, M. (2008). *Escaping North Korea: Defiance and hope in the world's most repressive country*. Lanham, MD: Rowan and Littlefield.

Kim, M. (2012). *Securitization of human rights: North Korean refugees in East Asia*. Santa Barbara, CA: Praeger.

Kim, S. (2013, July 31). Why are the majority of North Korean defectors female? *NK News*. Retrieved from https://www.nknews.org/2013/07/why-are-the-majority-of-north-korean-defectors-women/.

Kim, Y. (2009). *Long road home: Testimony of a North Korean camp survivor*. New York: Columbia University Press.

Kirkpatrick, M. (2014). *Escape from North Korea: The untold story of Asia's underground railroad*. New York: Encounter Books.

Ko, S. H., Chung, K. S., & Oh, Y. S. (2004). North Korean defectors: Their life and well being after defection. *Asian Perspective, 28*(2), 65–99.

Korea Institute for National Unification. (2016). *White paper on human rights in North Korea 2016*. Seoul: Center for North Korean Human Rights Studies. Retrieved from http://repo.kinu.or.kr/handle/2015.oak/7968.

Lankov, A. (2013). *The real North Korea: Life and politics in the failed Stalinist Utopia*. Oxford: Oxford University Press.

Lankov, A. (2015, August 14). Critique, but don't dismiss, N. Korean refugee testimony: Past treatment of Soviet, Cambodian atrocity survivors highlights dangers of disregarding defectors' stories. *NK News*. Retrieved from https://www.nknews.org/2015/08/critique-but-dont-dismiss-n-korean-refugee-testimony/.

Lankov, A. (2017). *Loyalists, defectors and life abroad: The daily life of North Koreans*. Wilmington, DE: NK Consulting Inc.

Lee, A. R. (2017). An investigation into the educational experiences of North Korean refugees in China in light of their exposure to human rights violations. *OMNES: The Journal of Multicultural Society*, *8*(1), 33–63.

Lee, H. (2015). *The girl with seven names: A North Korean defector's story*. London: William Collins.

Lee, K. (2006). *The border crossing North Koreans: Current situations and future prospects*. Studies series 06-05. Seoul: Korea Institute for National Unification.

Lee, S. (2016). *Every falling star: The true story of how I survived and escaped North Korea*. London: Amulet Books.

Lee, S. (2017, January 21). "*Every falling star*": Book talk by Sungju Lee. Presented at What the Book, Seoul, South Korea.

Lee, S. O. (1999). *Eyes of the tailless animals: Prison memoirs of a North Korean woman*. Bartlesville, OK: Living Sacrifice Book Company.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues, 2*(4), 34–46.

Liberty in North Korea. (n.d.). *Refugee rescues*. Retrieved from https://www.liberty-innorthkorea.org/rescue-refugees/.

McClintock, C. (2004). Scholar practitioner model. In A. Distefano, K. E. Rudestam, & R. J. Silverman (Eds.), *Encyclopedia of distributed learning* (pp. 393–397). Thousand Oaks, CA: Sage.

McEvoy, K., & McConnachie, K. (2012). Victimology in transitional justice: Victimhood, innocence and hierarchy. *European Journal of Criminology, 9*(5), 527–538.

Ministry of Unification. (2018). *Policy on North Korean defectors*. Retrieved from http://www.unikorea.go.kr/eng_unikorea/relations/statistics/defectors/.

Murphy, T. (2014). Authoritarian states. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 48–51). Thousand Oaks, CA: Sage.

NAUH. (n.d.). *Greetings from the President*. Retrieved from http://www.nauh.or.kr/eng/sub/sub01_02.php.

North Korea Strategy Center. (n.d.). *Vision and mission*. Retrieved from http://nksc.us/about-nksc/#ourteam.

Oberdorfer, D. (2001). *The two Koreas: A contemporary history*. New York: Basic Books.

Park, J. H. (2016, January 24). North Korean defector Yeonmi Park's public lecture in prosperity. *KnewsN: Online Korean News*. Retrieved from http://www.knewsn.com/kn/bbs/board.php?bo_table=KNEWS&wr_id=171.

Park, Y. (2015). *In order to live: A North Korean girl's journey to freedom*. New York: Penguin Books.

Patterson, M., Milburn, M., & Monteiro, C. (2014). Government propaganda. In T. R. Levine (Ed.), *Encyclopedia of deception* (pp. 418–422). Thousand Oaks, CA: Sage.

Reason, P., & Bradbury, H. (2008). *The SAGE handbook of action research: Participative inquiry and practice* (2nd ed.). Thousand Oaks, CA: Sage.

Robins, S. (2011). Towards victim-centred transitional justice: Understanding the needs of families of the disappeared in postconflict Nepal. *International Journal of Transitional Justice, 5*(1), 75–98.

Robins, S., & Wilson, E. (2015). Participatory methodologies with victims: An emancipatory approach to transitional justice research. *Canadian Journal of Law and Society, 30*(2), 219–236.

Robinson, M. E. (1982). Ideological schism in the Korean nationalist movement, 1923–1930: Cultural nationalism and the political critique. *Journal of Korean Studies, 4*, 241–268.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Simon and Schuster.

Sharf, B. F., & Vanderford, M. L. (2003). Illness narratives and the social construction of health. In T. L. Thompson, A. M. Dorsey, K. I. Miller, & R. Parrott (Eds.), *Handbook of health communication* (pp. 9–34). Mahwah, NJ: Lawrence Erlbaum.

Shin, M. D. (2018). *Korean national identity under Japanese colonial rule: Yi Gwangsu and the March First Movement of 1919*. New York: Routledge.

Son, S. A. (2015). Unity, division and ideational security on the Korean Peninsula: Challenges to overcoming the Korean conflict. *North Korean Review, 11*(2), 45–62.

Son, S. A. (2016). Identity, security and the nation: Understanding the South Korean response to North Korean defectors. *Asian Ethnicity, 17*(2), 171–184.

Son, S. A. (2018). North Korea's human rights insecurity: State image management in the post-UN COI Era. *Asia & the Pacific Policy Studies, 5*(1), 138–149.

Song, J. J., & Bell, M. (2018). North Korean secondary asylum in the UK. *Migration Studies.* https://doi.org/10.1093/migration/mnx074.

Sorrells, K. (2012). *Intercultural communication: Globalization and social justice.* Thousand Oaks, CA: Sage.

Starek, J. E., & Keating, C. F. (1991). Self-deception and its relationship to success in competition. *Basic and Applied Social Psychology, 12*(2), 145–155.

Stoller, P. (2016, June 12). The power of public scholarship. *Huffington Post*, Retrieved from https://www.huffingtonpost.com/paul-stoller/the-power-of-public-schol_b_10409082.html.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Thousand Oaks, CA: Sage.

Suh, Y. K. (2013). *Some problems of South Korean government's current integrating policies regarding North Korean defectors and its future options.* Retrieved from http://www.psa.ac.uk/sites/default/files/2852_637.pdf.

Tanaka, H. (2008). North Korea: Understanding migration to and from a closed country. *Migration Policy Institute*, Retrieved from https://www.migrationpolicy.org/article/north-korea-understanding-migration-and-closed-country/.

Transitional Justice Working Group. (forthcoming). *Survey on transitional justice for North Korea.*

Worrall, S. (2015, July 19). How a young girl escaped the prison that is North Korea. *National Geographic*, Retrieved from https://news.nationalgeographic.com/2015/07/150719-north-korea-refugees-children-famine-world-ngbook-talk/.

Index