

Performance Auditing

Performance Auditing

Contributing to Accountability in
Democratic Government

Edited by

Jeremy Lonsdale

Peter Wilkins

Tom Ling

Edward Elgar

Cheltenham, UK • Northampton, MA, USA

© Jeremy Lonsdale, Peter Wilkins and Tom Ling 2011

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2010939206



ISBN 978 1 84844 972 5

Typeset by Servis Filmsetting Ltd, Stockport, Cheshire
Printed and bound by MPG Books Group, UK

Contents

<i>List of contributors</i>	vii
<i>Foreword by Paul L. Posner</i>	xii

1. Introduction	1
<i>Jeremy Lonsdale</i>	
2. Performance auditing: audit or misnomer?	22
<i>Jan-Eric Furubo</i>	

PART I THE CONDUCT OF PERFORMANCE AUDIT

3. Selection of topics	51
<i>Vital Put and Rudi Turksema</i>	
4. Norms used: some strategic considerations from The Netherlands and the UK	75
<i>Vital Put</i>	
5. The right tools for the job? Methods, choice and context	95
<i>Jeremy Lonsdale</i>	
6. Evidence and argument	118
<i>Alex Scharaschkin</i>	
7. Forming judgements: how do VFM audit teams know what they know?	135
<i>Justin Keen</i>	
8. Standards and quality	147
<i>Peter Wilkins and Richard Boyle</i>	

PART II THE CONTRIBUTION OF PERFORMANCE AUDIT

9. The impact of performance audits: a review of the existing evidence	175
<i>Eddy Van Looke and Vital Put</i>	
10. Accountability, performance and performance auditing: reconciling the views of scholars and auditors	209
<i>Mark Funkhouser</i>	

11. On the effects, lack of effects and perverse effects of performance audit <i>Frans L. Leeuw</i>	231
12. Impact at local government level: a multiple case study <i>Katrien Weets</i>	248
13. Learning in an accountability setting <i>Jeremy Lonsdale and Elena Bechberger</i>	268
14. Responsiveness in performance auditing: towards the best of both worlds <i>Peter van der Knaap</i>	289
PART III CONCLUSIONS	
15. Conclusions: performance audit – an effective force in difficult times? <i>Jeremy Lonsdale, Tom Ling and Peter Wilkins</i>	311
<i>Index</i>	337

Contributors

Elena Bechberger is Head of Methods and Innovation at the National Audit Office (NAO) in London, UK. She studied public policy at the University of Konstanz (Germany) and holds a PhD in Government from the London School of Economics and Political Science (LSE). Previously she worked for the German Ministry of Economics and as an evaluator of European Commission projects, before joining the NAO as Research Fellow in 2006. She conducts research and training on qualitative and quantitative methods of programme evaluation, as well as on fiscal and social policy.

Richard Boyle is Head of Research, Publications and Corporate Relations with the Institute of Public Administration in Dublin, Ireland. He has worked with the IPA since 1986, having previously worked on policy and evaluation issues in local government in the UK. His research interests focus on public service modernization, managing for results in the public sector, and developing and implementing effective performance management and evaluation systems. He has researched and written extensively on public service reform, performance measurement and evaluation. He has worked widely with Irish central and local government, the OECD and the World Bank. He was a member of the Board of the European Evaluation Society from 2002 to 2005 and is chair of the Irish Evaluation Network.

Mark Funkhouser is currently the Mayor of Kansas City, Missouri, USA. He holds a BA in Political Science, an MSW, an MBA, and a PhD in Public Administration and Sociology. He authored the book *Honest, Competent Government: The Promise of Performance Auditing* (2009), which shows how auditing can increase citizens' trust in government. Mark has also authored numerous articles and essays that have been published in journals such as *Public Administration Review*, and in books including *Public Administration: Cases in Managerial Role-Playing*. In addition, Mark was the editor for *Local Government Auditing Quarterly* for 9 years. Throughout his 40-year career he has been a vocational rehabilitation counselor, an assistant professor, and a city auditor. In 2003 Mark was named one of *Governing Magazine's* 'Public Officials of the Year'.

Jan-Eric Furubo works at the Swedish National Audit Office, Stockholm, where he is responsible for focusing the long-term development of performance audit. He has published widely on evaluation methodology, the role of evaluation in democratic decision-making processes and its relation to budgeting and auditing. He co-edited the *International Atlas of Evaluation* (2002) and *Evaluation – Seeking Truth or Power?* (2010). He contributed to *Handbook of Public Policy Analysis – Theory, Politics and Methods* (2007), was co-author of an article about Evaluation Systems published in *Evaluation* (2008) and contributed to *Evaluer les politiques publiques pour améliorer l'action publique, Une perspective internationale* (2009).

Justin Keen was the first NAO Fellow in Audit and Evaluation, 1996–98. Following his time at the NAO he worked at the King's Fund, a London-based charity, and at University College London, UK, before moving to his present post as Professor of Health Politics at the University of Leeds in 2001. He has undertaken a number of evaluations of health policies over the last 20 years, and has a particular interest in the development of more robust methods for policy analysis.

Frans L. Leeuw is the Director of the Justice Research, Statistics and Information Centre, affiliated to the Dutch Ministry of Security and Justice, and is Professor of Law, Public Policy and Social Science Research at Maastricht University, the Netherlands. He is a sociologist (PhD Leyden University, 1983). Earlier he was Professor of Evaluation Studies at Utrecht University, Director of the Performance Auditing and Evaluation Department of the Dutch National Audit Office, Dean of the Humanities and Social Sciences Faculty of the Netherlands Open University, Chief Inspector for Higher Education in the Netherlands and Associate Professor of Policy Studies at Leyden University. He was one of the founders of the European Evaluation Society and past president of this body. Currently, he is President of the Dutch Evaluation Society. He is Faculty Member of the IPDET program. He has worked for the World Bank, the EU and many agencies and ministries of the Dutch government.

Tom Ling is Director for Evaluation and Performance Audit at RAND Europe, in Cambridge, UK. He studied Social and Political Sciences at Cambridge University, UK, and completed a PhD in Government at Essex University, UK. He joined RAND following 4 years as Research Fellow at the National Audit Office in the UK. He has worked on evaluation projects with the European Commission, UK Government departments, the National Audit Office, the Health Foundation and many

others. He has published widely on evaluation, accountability and related topics. He recently completed a Performance Audit handbook for RAND Europe, and co-edited *The Evidence Book*, a critical examination of the use of evidence in public policy and service delivery (published by Transaction in 2010). He is Professor of Public Policy (Emeritus) at Anglia Ruskin University, UK, and honorary senior visiting research fellow, University of Cambridge, UK.

Jeremy Lonsdale is Director-General, Value for Money Audit at the National Audit Office (NAO) in London, UK. He studied history at King's College, University of London, UK, is a qualified accountant and has a PhD from Brunel University, UK, supported by the NAO in 2000. He was co-author of *Performance or Compliance? Performance Audit and Public Management in Five Countries* (OUP, 1999) and co-editor of *Making Accountability Work: Dilemmas for Evaluation and for Audit* (Transaction, 2007). His most recent publication was 'The Audit Society: Helping to Develop or Undermine Trust in Government' (written with Irvine Lapsley) in *Administrative Justice in Context* (Hart, 2010), edited by Michael Adler.

Vital Put has worked – from 1998 onwards – at the Belgian Court of Audit, Brussels, first as a performance auditor, later as a performance audit supervisor. In addition, since 2003, he has worked as a researcher at the Public Management Institute of the University of Leuven. He previously worked at several Belgian Ministries (Justice, Labour, and Social Affairs). Vital possesses a master degree in public management and in criminology and a PhD in social sciences. He has published several articles and chapters on performance auditing.

Alex Scharaschkin is Director for Consumers and Competition Value-for-Money studies at the National Audit Office in London, UK. He currently leads the NAO's value-for-money work on economic regulation and the use of market mechanisms in public service delivery, and has previously led value-for-money studies in the health area. He is interested in the use of statistical and economic methods in value-for-money assessment, and in the nature of evidential support for conclusions on value-for-money more generally. He joined the NAO in 2000, having previously held posts in the Civil Service and the education sector, and at the Universities of London and Melbourne. His academic background is in mathematics, but he has also carried out and published research in applied psychological and educational measurement.

Rudi Turksema is a Performance Audit Expert at the Netherlands Court of Audit, The Hague. His work focuses on the quality and effectiveness of

performance audits. This also involves the quality of performance information, evaluation studies, and results-based budgeting. Before coming to the Netherlands Court of Audit, he received his PhD in Sociology (on a study into the effect of government policy on day-care supply) from Utrecht University, the Netherlands.

Peter van der Knaap is Performance Audit Director with the Netherlands Court of Audit in The Hague. Prior to this, he was Head of Policy Evaluation at the Ministry of Finance in The Hague, where the introduction of the performance-oriented budgeting system and corresponding evaluation framework were his main responsibilities. Before that he worked at the Inspectorate of the Budget and was at Erasmus University, Rotterdam, where he earned his PhD in 1997. He publishes and lectures on policy audit and evaluation, performance-based budgeting and policy-oriented learning.

Eddy Van Loocke entered the Belgian Court of Audit in Brussels as a junior auditor in 1985. When the Court was charged with the legal mandate to carry out performance audits in 1998 he focused his activities on this new mandate. He has been an audit manager since 2004 and has conducted many performance audits and participated frequently in peer reviews and audit supervising committees. He coaches audit teams in the field of planning and conducting performance audits. He is also a member of the coordination committee of the Flemish Evaluation Society.

Katrien Weets has been working as a researcher at the Public Management Institute of the Katholieke Universiteit Leuven in Belgium since 2007. She holds a master's degree in political sciences (public management). As a researcher associated with the Policy Research Centre – Governmental organization in Flanders, she conducts both fundamental and practice-oriented research. Her main research interests include public performance management, public budgeting and public auditing. Currently, Katrien is working on a PhD on municipal planning and budgeting in Flanders.

Peter Wilkins is the Deputy Ombudsman of Western Australia and was for several years Assistant Auditor General, Performance Review for the Western Australian Auditor General. He has extensive practical and research experience regarding public sector performance review, reporting and accountability. He liaises regularly with ombudsmen and audit institutions in Australasia, North America and Europe and has contributed to national and international initiatives. He also holds the positions of Adjunct Professor in the School of Accounting and The John Curtin Institute of Public Policy at Curtin University of Technology, Western

Australia, Associate of the Public Sector Governance and Accountability Research Centre at La Trobe University, Victoria, Australia and is a member of the International Evaluation Research Group.

Foreword

Paul L. Posner

Accountability has always been central to the practice of public administration in any democratic system. As the principal mechanisms through which the people control their leaders, the field of public administration has rightly focused considerable attention on the concept of accountability.

In many OECD nations, the term accountability has grown to an iconic status, with a symbolic imagery that permits this chameleon-like term to be attached to a wide range of causes and agendas.¹ Rising expectations for public provisions of services have been accompanied by growing pressures for accountability on the part of the proliferating list of stakeholders and participants in the policymaking process. Public organizations are condemned to live in a world where the stakes associated with public programs have grown, but where there is little agreement among the many players on goals, expectations or standards.

The rather straightforward, simple world of early public administration has been replaced by a world where public leaders and administrators must attempt to meet conflicting accountability expectations. Indeed, modern treatments consider accountability to be a multi-faceted concept, encompassing separate and often competing accountabilities to internal and external stakeholders. One classic treatment suggests that most government agencies are simultaneously beholden to four systems of accountability: (1) bureaucratic, (2) legal, (3) professional and (4) political. It goes without saying that each of these perspectives can and does conflict, leaving public managers with the difficult job of balancing these differing obligations.² Most public organizations are, in effect, agents for multiple principals both within and outside their boundaries. Some have gone so far to suggest that most public organizations suffer from Multiple Accountabilities Disorder!³

Most advanced nations have responded to accountability imperatives by articulating institutional reforms focused on the performance and results of government operations and programs. Performance measurement and policy analysis have become mainstreamed into management

and budgeting over the past several decades. Many policy advocates and analytic organizations outside of government work tirelessly to transmit analysis and information to policymakers, either putting new issues on the table or providing oversight and insight on existing programs or operations.

However, it is the transformation of audit institutions that has been most notable. National audit offices have expanded their remit beyond traditional financial and compliance auditing to focus on performance auditing and assessments. Performance auditing has become a central feature of most advanced nations' national audit offices, and this has often been replicated, in the United States, in inspectors' general offices located inside agencies. Indeed, some offices have been pushed into ever more expansive policy roles, becoming authenticators of new problems pushing their way onto policy agendas and adjudicators of budget forecasts in addition to their traditional program review responsibilities.

I witnessed these trends from my position at the United States Government Accountability Office (GAO), where I served for many years as a managing director responsible for GAO's federal budgeting work. An audit agency that began in the 1920s, with the responsibility of reviewing all financial transactions by federal agencies with thousands of financial auditors, it has become a much smaller agency with a multi-disciplinary workforce devoted to performance and program assessments. The inspectors general in federal agencies, as well as state and local auditors, have followed as political leaders and restive publics place more demanding and complex tasks on government than ever before. Federal auditors are now asked by the Congress and the President to go beyond even post-performance audits to become more proactive in working with managers to mitigate and prevent potential waste and fraud when programs are started.

As they have expanded their roles to performance, auditors have succeeded in achieving major influence in forming the policy agenda and in formulating public policy as well. In the United States, GAO input was critical in bringing about reforms in policy areas ranging from healthcare reimbursement formulas, grant allocations, and reforms of federal deposit and pension insurance programs. In one prominent case in Canada, the Auditor General's reports on the Martin government's pattern of influence peddling was widely viewed as the most important event triggering the government's downfall.

However, auditors achieve their influence in highly contestable systems, rife with competing values. Far from hegemonic influence, these systems appear to veer from accountability deficits to accountability excesses, depending on such variables as the strength of accountability offices and

the receptivity of the broader political system. Auditors often have sufficient credibility to make powerful claims on the agenda, often prompting government officials to modify their agendas to address the issues raised in reports. In some parliamentary systems, the auditor general is provided with formal opportunities to testify before oversight committees chaired by the minority party. But other actors are competing for influence as well, including established interest groups and bureaucratic agencies, who can prove to be formidable contestants in protecting their programs and claims.

We know surprisingly little about what conditions aid or hinder the influence of audit offices in the policy process. One obvious one is the orientation and capacity of the audit office itself. Specific features of accountability institutions themselves limit their role in policymaking. For example, traditionally, many audit agencies have not engaged in policy advocacy. They have been closed organizations concerned with their independence and reluctant to work with others. They have not, therefore, looked to form coalitions. Such closed organizations have kept their draft reports to themselves for fear of leaks and have privileged access to information that they cannot share. And they generally chose not to speak publicly on issues or to get involved with others. Many auditing organizations rotate their staff when they become too familiar with the programs they are reviewing, which preserves independence, but often at the expense of expertise. This institutional insulation and isolation has the price of diminishing the roles such institutions can play in the issue networks that are responsible for policy development and change in most systems.⁴

Those audit institutions that are more fully engaged with policymakers face institutional risks of their own. Most of the work done by the GAO is at the request of members of Congress from both parties. The agency must delicately steer between responding to these legitimate information needs to set their research agenda while sustaining their independence in developing findings and reports. The broader engagement of these institutions with media, universities and other actors also can sustain their reputations which can promote support and limit interference. However, the involvement of accountability professions in reviewing program results and effectiveness carries obvious political risks for audit institutions – many have charters that limit their coverage of these issues and constrain them from making recommendations on policy and program design issues.

Notwithstanding the greater centrality of auditors in public management, public managers and policymakers often have little connection or dialogue with these independent officials who are the institutional champions of accountability. Managers on the front lines of program performance often have no familiarity with audit standards, materiality criteria

or any of the other tools that are so central to the job of the performance auditor. The vaunted independence that is so essential to the credibility of auditors also has served to wall off these accountability officials from their managerial counterparts in the agencies of government.

Several years ago at an international meeting, I succeeded in bringing together the senior budget directors of several major OECD nations with the heads of the national audit offices of those nations. What was remarkable was how little they knew about each other; indeed one budget director remarked that he came to the meeting to finally get a chance to get to know how this strange and mysterious institution, which had become so important to governance in his and other nations, did its work.

The separation between managers and auditors in the world of practice has its parallel in the academic community. While accountability concepts have been a foundation of public administration over the years, there has been precious little focus on audit institutions. For instance, the *Public Administration Review*, among the premier journals on the field in the United States and the world, has only one article with audit institutions in its title in the past ten years, and only six articles with some coverage of those institutions.⁵ One would have to go to accounting and auditing journals to find research on the role and management of audit institutions in public administration.

This book is a much welcome tonic for public administration. It is one of the few books that explicitly focus on how audit institutions carry out their performance auditing responsibilities. While auditors will likely read it, the authors have geared the book to a broader readership, including public managers who are often the subject of performance audits.

It is also notable that the book has contributions from both practitioners as well as academics who cover audit institutions. This is increasingly rare but extraordinarily valuable. Many of the authors are, in fact, ‘pracademics’ who have one foot in public office and one in academia teaching or doing research. The giants who founded our field of public administration in the United States – Woodrow Wilson, Luther Gulick, Charles Merriam – were themselves all pracademics who enriched their theory with illustrious service in government, and vice versa.⁶ Readers will benefit from the authors’ mix of first hand experience and reflective scholarship – both essential for a deep and rich understanding of developments in our field.

The chapters in the book are notable for their coverage of important issues. There are chapters covering issues ranging from organizational strategy, audit tools and methods, and standards. These chapters offer a view into the operations of audits that many public managers know precious little about.

The chapters on the impacts of performance audits are very revealing.

They help us gain a better understanding about the roles that performance audits play in the broader policy process. Auditors themselves need more sophisticated analysis of how they achieve the impacts on policymaking and this book will provide a rich source to help them learn about the intended and unintended effects of their work. The comparative focus of the book is also a contribution, with chapters including material on audit bodies in Belgium, the Netherlands, the United Kingdom, Australia and the United States, among others.

My hope is that this book will help trigger an expanded dialogue between auditors, public managers and students of public administration. While audit reports rightly criticize the stovepiping of government agencies and programs, our own field has erected its own professional stovepipes that have inhibited networking and valuable collaborations across the differing disciplines that are engaged in public management. Constructive engagement can promote the objectives of both managers and auditors without jeopardizing the independence that is so essential to the credibility of audit institutions. As a firm believer in the important role of ideas in public policy, I have faith that books like this one can help to bring about greater integration of theory and practice of public administration in the future.

Paul Posner is professor and director of the Masters in Public Administration program at George Mason University. He was previously Director of the Federal Budget and Intergovernmental Relations at the Government Accountability Office in Washington DC. He has a PhD from Columbia University.

NOTES

1. Melvin J. Dubnick and Jonathan B. Justice (2004), 'Accounting for Accountability', paper delivered at the Annual Meeting of the American Political Science Association, September.
2. Barbara S. Romzek and Melvin J. Dubnick (1987), 'Accountability in the Public Sector: Lessons from the Challenger Tragedy', in *Public Administration Review* 47 (May/June).
3. Thomas Schillemans and Mark Bovens (2008), 'The Challenge of Multiple Accountability: An Empirical Analysis', paper prepared for Kettering Symposium on Public Accountability, Dayton, Ohio, May 22–4.
4. Rob Schwartz and Paul L. Posner (2008), 'Accountability Institutions and Information in the Policy Process', paper prepared for Kettering Symposium on Public Accountability, Dayton, Ohio, May 22–4.
5. Search by the author in Blackwell Synergy index of *Public Administration Review*, 2000–10.
6. Paul L. Posner (2009), 'The Pracademic: An Agenda For Re-engaging Practitioners and Academics', *Journal of Public Budgeting and Finance* 29:1 (March).

1. Introduction

Jeremy Lonsdale

Performance audit involves assessing whether government policies, programs and institutions are well managed and being run economically, efficiently and effectively. This is a task of potentially great significance – at a practical level for citizens, and at a more abstract level for the health and vitality of democratic governance. (Pollitt et al., 1999)

The post-bubble era is going to be impatient of extravagance. It will be flintier, value-conscious and much less forgiving of waste. It will demand that the public sector justifies its existence to those who pay its bills . . . Simmering jealousy will boil up into hot anger if the public sector isn't delivering value for money. (Andrew Rawnsley, *The Observer* newspaper column (UK), 2009)

The first decade of the twenty-first century will be remembered for the scale of the economic crisis experienced across the world, and the lives of individuals in the current decade will be shaped by how governments respond to the fundamental changes that have occurred. Total capitalisation of the world's stock markets was almost halved in 2008, representing a loss of nearly US \$30 trillion of wealth. Industrial production in the first quarter of 2009 fell 62 per cent in Japan and 42 per cent in Germany.

The global crisis forced governments to act swiftly under considerable pressure in order to tackle a range of political and economic problems affecting both the private and the public sectors. These actions included immediate intervention in the financial sector, massive budgetary commitment to extra-ordinary public spending, quick expansion of public sector capacity to deal with the effects of the crisis such as a rise in unemployment, and reprioritisation of existing spending programmes to provide more substantial economic effect from government activity. The most substantial activity has been seen in the United States, where an estimated \$800 billion stimulus package under The American Recovery and Reinvestment Act has been put in place, but similar measures have been seen in many other countries.

The scale and impact of the crisis have also required governments to focus their attention on potential wasteful spending in the public sector, and public bodies have come under considerable pressure. This pressure is to be seen in the form of:

- significant reductions in public employment;
- major cutbacks in programme spending, especially in areas of discretionary activity;
- repeated efficiency savings programmes;
- efforts to reduce regulation within government; and
- increased attention to securing the benefits from ICT initiatives, process re-engineering, mergers of bodies, shared services and market-type mechanisms.

Deficit reduction plans, prompted by the economic crisis, have provided governments with a new opportunity and a pressing need to focus on securing value for money from public spending. In the United Kingdom, the 2009 Labour Budget promised austerity in the public sector until at least 2017–18, and the Conservative-Liberal Democrat government, elected in May 2010, made immediate announcements of cuts in public spending as part of efforts to reduce the £156 billion budget deficit. Cost-cutting and efficiency programmes have been introduced by government bodies across the world, along with efforts to reduce procurement costs and increase public sector productivity. Some bodies have reduced the range of services they offer or changed eligibility rules. At the time of writing, it was still not clear whether efforts to reduce spending will lead to a ‘slash and burn’ approach in some countries, or whether cost reductions will be more carefully considered.

Efforts to increase the performance of the public sector are likely to be more dramatic in some countries than others. In the UK and US, for example, the large bank bailouts are seen as necessitating radical changes. In other countries less affected by the banking crisis, the public sectors are not expected to face such significant challenges. Nevertheless, declining tax revenues, the cost of renewing infrastructure, unfunded public-sector pensions and the implications of an ageing population – all issues which, along with the environment, were facing governments before the economic crisis – are all driving governments to turn their attention to the ways in which public services are delivered.

The events of 2008–10 have increased enormously the significance of concerns about public expenditure and value for money, but long before the current crisis, governments around the world had been developing measures to improve the management of the public sector and secure savings and improvements in performance. Talbot (2010) notes that ‘Performance measurement and management of public services has been on the rise in many countries in recent years.’ Bouckaert and Halligan (2008) describe performance as one of the two big questions in public management of the last 15 years (the other being the role of markets).

Steps have also been taken to improve the assessment of public sector performance and extend the information available to the public in order, it has been argued, to enhance accountability and transparency. In 1993, in the United States, the Government Performance and Results Act was introduced, ‘designed to improve the effectiveness of federal programs and citizen satisfaction through the systematic measurement and reporting of performance measures that are aligned to agencies’ goals’ (De Lancer Julnes, 2006). In 2000, the Canadian government introduced Results for Canadians, a management framework covering results, spending and value-for-money issues. Around the same time, the Swedish government introduced a range of initiatives on performance budgeting, and in Australia a performance management framework was put in place in 2001. In the UK, there have been repeated efforts to improve efficiency. The Gershon Review was carried out in 2004 which identified scope for £20 billion worth of savings. The Comprehensive Spending Review in 2007 referred to the government’s ‘far-reaching value-for-money programme . . . releasing the resources needed to sustain the pace of public service improvement and meet the global challenges of the decades ahead’ (Treasury, 2007). In 2008, the government launched what it described as two value-for-money programmes – an Operational Efficiency programme and the Public Value Programme. In the 2009 Budget, the UK government spoke of the ‘constant effort to improve value for money’ and issued a Value for Money Update to outline what departments had achieved (Treasury, 2009).

This short summary of recent events highlights three key aspects of government which run through this book. These are concerns for:

- the value to be achieved from public expenditure;
- the mechanisms for accountability for public sector performance in the face of changing forms of governance; and
- the credibility and trustworthiness of government and the information provided by it.

The events of the last few years have increased the importance of these issues, rather than introduced them afresh to political debate. But these developments – and their ongoing consequences – have only added to the significance of all forms of performance management, and encouraged debate as to whether the actions they generate to improve public sector performance are up to the job. Looking for an answer to that question as it relates to one form of performance management – performance audit – is the justification behind this book.

PERFORMANCE AUDIT AND THE PURSUIT OF ENHANCED PUBLIC SECTOR PERFORMANCE

The array of performance management and measurement techniques introduced in the last 30 years have included – in no particular order – contractual arrangements, targets, inspection regimes, customer charters, reorganisations, the creation of ‘arm’s length bodies’, performance-related pay, purchaser–provider splits, league tables, compulsory competitive tendering, and forms of consumer choice. All these measures have required the collection of data, in often intrusive ways. Alongside these developments has been the expansion of different forms of audit (Power, 1997, Lapsley and Lonsdale, 2010). Audit in its many forms has been seen by governments as an answer to many problems, inevitably those associated with the use of public money, but also increasingly the wider monitoring of public sector performance. Traditional audit – the scrutiny of financial statements – has continued and developed in complexity, but has been joined by other forms, including environmental audit, management audit, forensic audit, data audit, intellectual property audit, medical audit, and what is widely known as performance audit, the topic of this book.

Performance audit has developed over several decades in many parts of the world with the aim of assessing aspects of how government organisations have performed and have used the resources provided to them. It grew initially in Europe, Australasia and North America during the 1970s and 1980s, taking root in many democracies – at national level within Supreme (or state) Audit Institutions (SAIs), and variously at state, provincial, regional, local and municipal levels – as auditors have widened their perspective, from whether the money has been spent as intended, to broader considerations of whether it has been spent efficiently and, even more challengingly, whether it has been used to good effect (see Table 1.1).

In Sweden, for example, performance auditing became the major element of the work of the former national audit office, the Riksrevisionverket (RRV), when it was restructured in 1967 and financial audit was removed from its remit. Thereafter, the RRV was at the forefront of performance audit developments internationally in the 1980s and 1990s. In Finland, dissatisfaction with solely legalistic audit grew in the 1970s and performance audit was established as a separate form of audit in 1988. In the UK, ‘value for money’ audit was given legal standing in 1983 following several decades of examinations of expenditure and revenue generated by the audits of the annual accounts. By the end of the 1990s the NAO was publishing 50 major performance audit reports a year (Pollitt et al., 1999).

Table 1.1 Performance audit mandates in selected countries

Year	Country	Legislation
1921	United States of America	General Accounting Office established with broad mandate to investigate 'all matters relating to the receipt, disbursement, and application of public funds' and 'to make recommendations looking to greater economy and efficiency in public expenditures'. Subsequent acts have clarified and expanded the mandate
1948	Austria	Federal Law on the Rechnungshof enabled it to examine the economy, efficiency and effectiveness of the operations of corporate public bodies, local authorities and provincial governments, and the economy and efficiency of state economic enterprises
1967	France	Legislation provided for the SAI, the Cour des Comptes, to examine aspects of the economy, efficiency and effectiveness of public money
1976	Netherlands	Government Accounts Act broadened out remit to performance audits to determine performance of government, organisation and management services. It was extended in 1992 to allow for examination of policy
1977	Canada	The Auditor General Act 1977 provides the original legal basis for the Auditor General to carry out performance audits. It was amended in 1995 to include responsibilities related to environmental matters
1983	United Kingdom	National Audit Act formalised the NAO's ability to examine the economy, efficiency and effectiveness of government spending
1993	Ireland	The Comptroller and Auditor General (Amendment) Act allows the C&AG to carry out examinations of the extent to which acquisition, use and disposal of resources have been carried out economically and efficiently, but not directly look at effectiveness
1997	Australia	Auditor General Act authorises the Auditor General to conduct a performance audit of an entity, a Commonwealth authority or company, other than a Government Business Enterprise
1998	Belgium	Provides for the audit of the sound use of public funds and to examine economy, efficiency and effectiveness

Performance audit evolved with the times, taking on new forms. In the Dutch audit office, the Algemene Rekenkamer, for example, audit units were set up in the 1960s to investigate the efficiency of government organisations and early computer systems. In 1976 the Government Accounts Act broadened the remit to carrying out performance audits to determine the performance of management, organisation and government services. In the 1980s different types of performance audit were developed, many focusing on providing assurance that plans were being implemented, and others taking the form of benchmarking to compare ministries. In 1992, the Government Accounts Act was amended to allow the Rekenkamer to examine policy and in the 1990s it examined the consequences of policy on a limited scale. This grew after 2001 into a statutory task to examine the effectiveness and efficiency of policy. The focus has been on whether the policy has the desired effect, focusing on social problems, and trying to identify whether there is a gap between policy and implementation (Algemene Rekenkamer, 1989, 2010).

More recently we have seen the development of performance audit, for example, in eastern Europe, Asia (for example, in Japan and South Korea) and Latin America. In Africa, performance audit has been undertaken in Botswana for nearly 20 years and was first carried out by the Auditor General of South Africa in 1986. AFROSAI-E, the African Organisation of English Speaking SAIs, now describes performance audit as ‘one of its six imperatives’ (AFROSAI, 2009) and argues that to comply with international standards and recommendations it needs to increase ten-fold the number of performance auditors (currently, 300). In 2009, a competition for the best performance audit from an African SAI was won by the Ghanaian Audit Office. Some of this work has been taken forward with western European SAI assistance – for example, training courses provided by the Swedish NAO since the 1980s, or capacity building and input into studies as provided by the Canadian Office of the Auditor General and the UK NAO to the Chamber of Accounts in the Russian Federation. It is also supported by international bodies such as the World Bank, or the Asian Development Bank, for example, through its work to enhance performance audit capacity in China.

Definitions and Characteristics

The focus of this book is primarily on SAIs working at national level and so to seek a definition of performance audit it is appropriate to start with INTOSAI, the ‘umbrella’ organisation that acts as a forum for the SAIs around the world. The INTOSAI Auditing Standards (INTOSAI, 2004) state:

Performance auditing is concerned with the audit of economy, efficiency and effectiveness and embraces:

- (a) audit of the economy of administrative activities in accordance with sound administrative principles and practices, and management policies;
- (b) audit of the efficiency of utilisation of human, financial and other resources, including examination of information systems, performance measures and monitoring arrangements, and procedures followed by audited entities for remedying identified deficiencies; and
- (c) audit of the effectiveness of performance in relation to achievement of the objectives of the audited entity, and audit of the actual impact of activities compared with the intend impact.

Within this broad definition, performance audit has developed different forms. Some writers have distinguished between substantive performance audits, and systems and procedures performance audit (Shand and Anand, 1996). The former tend to consider efficiency and effectiveness issues, whilst the latter focus on the systems and procedures used to deliver and evaluate programmes. Another analysis (Bowerman, 1996) distinguished between reviews of management systems, arrangements and procedures, the performance procedure audit, policy audit, audit of management representations of performance, comparative performance audit, and quality audit. More recently, Grönlund et al. (2009) in a review of reports by the Swedish NAO (SNAO or Riksrevisionen) published between 2003 and 2008, identified eight types of performance audits – covering economy, efficiency and effectiveness, systems, administration, goal-related, policy and empirically grounded audits.

Different interpretations have led to work with different focus. For example, in two Australian state audit bodies – Western Australia and Tasmania – performance auditing encompasses ‘the range of audit and review activities from annual attest work on financial statements and performance indicators through to the preparation of direct reports on performance examinations’ (Nichol, 2007). In Canada, the Office of the Auditor General states that its performance audits ‘examine the government’s management practices, controls and reporting systems with a focus on results’ (OAG, 2010). In the United Kingdom, the National Audit Office’s form of performance audit is designed to gather evidence so as to conclude on whether ‘value for money’ has been achieved, a term it defines as ‘the optimal use of resources to achieve the intended outcomes.’

SAIs have developed their performance audit to meet the needs of their specific environments. In Sweden, for example, the SNAO states ‘Performance audits should primarily concentrate on circumstances related to the government budget, or to the implementation and

results of government activities and undertaking. They may also refer to government activities in general' (SNAO, 2008, 12). It adds that performance audits must 'concentrate on issues that are important to society and in which there are clear risks of shortcomings in efficiency' and states that the selection of audits must be based on the following criteria:

- The audit should be based on a presumed problem.
- The audit should concern actual government activities that are either ongoing or implemented.
- The basis of assessment for the audit must be clearly stated. It should normally either stem from direct decisions by the Riksdag (or in certain cases the government) or be possible to infer from the direct or indirect standpoint of the Riksdag.
- It should be possible to answer the audit questions with a high degree of reliability and precision.
- Reports by the SNAO should be drawn up in such a way that they may provide a basis for demanding accountability. It is therefore important in the audit to be able to apportion responsibilities and criticism between the parties involved.

SAIs have also taken their work forward differently through choice or through their interpretation of what is needed. As mentioned earlier, the Algemene Rekenkamer in the Netherlands, for example, has chosen to focus its work on the gap between policy intentions and implementation because, based on past performance audits, it considers that policy formulation is over-valued and policy implementation under-valued, leading to poor value for money for the Dutch taxpayer. It states that its performance audit work:

investigates whether the ministers' policies are effective: do they produce the intended results? We look at whether the intended results can be achieved by means of the ministers' policies. We also consider the implementation of policy: does the minister do what has been agreed and are the results for society? (Algemene Rekenkamer, 2009)

In part, such variety in the nature of the work is a reflection of the different regulations governing performance audit within the different jurisdictions. Some SAIs, for example, do not have the remit to examine effectiveness. In Australia, one of the eight states and territories excludes reviews of management and agency effectiveness, whilst two focus on systems rather than management (Nichol, 2007). In Canada, the legislation excludes assessing effectiveness directly, whilst in Ireland the legislation allows the

Comptroller and Auditor General (C&AG) to examine the systems, procedures and practices employed by bodies to evaluate the effectiveness of operations, making comparisons with other bodies, practices and systems as he sees fit, but not examine the effectiveness of activities directly himself.

The skills and disciplines of performance auditors also differ, helping to shape the work. In some jurisdictions the work is undertaken by professional accountants, in others it can be conducted by lawyers, or those with backgrounds in public administration, social or political sciences or some other discipline. In 2005, more than 80 per cent of European SAIs reported employing staff with legal qualifications, and all but the French Cour des Comptes employed staff with accountancy or business qualifications (National Audit Office, 2005). Different skills lead to different approaches. In Norway, for example, the introduction of new staff led to a more evaluative approach to performance audit (Gunvaldsen and Karlsen, 1999), whilst Lonsdale (2008 and in this book) draws attention to the methodological developments at the UK NAO which have arisen, for example, from bringing in staff with different skills or by making use of outside experts to help carry out audits.

The institutional settings within which performance audit is conducted also shape the work. They all have a place in the system of government provided for by the constitution or statute and have similar key characteristics – independence from government, professionally qualified staff, strong powers of access to information and documents, and the ability to report freely. But within Europe, for example, there are four distinct models of SAI (National Audit Office, 2005) – the court with a judicial function such as the French Cour des Comptes, a collegiate structure with no judicial function as can be seen in the Netherlands and Germany, an audit office headed by an auditor-general or president such as in the UK, and a ‘distinct model’ headed by a president and auditing at central, regional and local level, as in Austria. Considering how constitutional positions affect the way in which work is conducted in France and the UK, Astill and Page (2006) commented:

the judicial position of the Cour and the fact that Cour, like the French judiciary, is a state institution staffed by civil servants, means that the approach to these questions of authority, trust and expertise can be characterised as an ‘insider’ approach. It uses the authority, trust and expertise that are supposed to reside in a state institution to achieve the impact on the bodies it audits. The NAO . . . pursues what can be classed as an ‘outsider’ strategy: establishing its impact through proving its worth to those inside government through consultation with them and relying on groups and forces outside Whitehall to develop its conclusions and also nudge government to accept its conclusions when they are reluctant to do so.

BRIEF OVERVIEW OF SEVERAL NATIONAL PERFORMANCE AUDIT REGIMES

This book cannot cover the development of performance audit in all its variety throughout the world. Instead, our focus is on aspects of the conduct of the work in countries where performance audit has a significant, well-established tradition: Australia, Belgium, the Netherlands, Sweden, the United Kingdom and the United States (Table 1.2). The work takes many different forms and is used in a variety of ways by legislatures and governments.

In Australia, performance audits are undertaken by the Australian National Audit Office (ANAO) and by the state audit bodies. In 2008–09, the ANAO focused on themes of governance and project management; border security and national security; community support and well-being; environment; industry, science and education; and transformation of entities. Forty-five performance audit reports were prepared at an average cost of A\$0.39 million, with topics ranging from the Management of Funding for Public Works, Army Reserve Forces, Centrelink's Complaints Handling System, and Tourism Australia. The ANAO's reports are reviewed by the Joint Committee of Public Accounts and Audit. The Committee reviewed 26 performance audits in 2008, supporting all ANAO recommendations. Its reports detailed the findings of the committee's examination of the audit reports during that parliament and chose five audit reports for particular scrutiny (ANAO, 2009).

The Belgian Court of Audit undertakes audits at the federal level as well as at the regional and provincial level. In 2009–10, its performance audits covered topics such as the organisation and working of the Special Tax Inspectorate, the knowledge and conservation of collections by the Royal museums of fine arts and the Royal museums of art and history, the implementation of the Kyoto protocol, the use of scientific knowledge in healthcare policy, public–private partnerships, support for people with difficulties integrating into the labour market, staff planning in government, rational use of energy in buildings, the functioning of the office for employment and professional training, and educational and administrative support given to elementary and secondary schools. The reports of the Belgian Court of Audit are discussed by the committees in the legislatures to which they are submitted – the Federal Parliament, the Flemish Parliament and the Walloon Parliament.

In the Netherlands, the Algemene Rekenkamer investigates whether central government revenue and expenditure are received and spent correctly, and whether central government policy is implemented as intended. As part of this remit, its performance audits investigate

Table 1.2 *The role of performance audit*

	Use made of work	Resources employed	Outputs
Australian National Audit Office	Reports taken by Joint Committee of Public Accounts and Audit	\$23.594 million in 2008–09	45 reports in 2008–09
Rekenhof (Belgium)	Used by policy committees in parliaments	Use of external experts in exceptional cases	15–20 reports per annum
Algemene Rekenkamer (Netherlands)	Used by parliamentary committees, and also government bodies	80 staff, although some part-time on performance audit 3.9m EURO (21% of audit costs) Occasional use of external experts	15 per annum
Riksrevisionen (Sweden)	Used by various parliamentary committees, government ministries and agencies	100 staff 133 million Swedish Crown (46% of budget) Around 20% of studies used considerable number of external experts	33 reports in both 2008 and 2009
Government Accountability Office (USA)	Reports directly to Congress. During 2007, 92% of its performance audits were undertaken in response to a specific congressional request or statutory mandate	Draws on its total staff of 3,100 Limited use made of external resources	Published 1,038 performance-audit products in 2007
National Audit Office (UK)	Used by House of Commons Committee of Public Accounts as basis for scrutiny hearings	Around 270 staff including trainees Considerable use of external experts	63 reports in 2009–10

whether ministerial policies are effective and produce the intended results (Algemene Rekenkamer, 2010). It states that it audits ministerial efficiency to see whether the right amount of money is used to achieve the intended results. It focuses on central government but also undertakes performance audits of institutions that use public funds to carry out, at arm's length, statutory tasks. These bodies include schools, benefit agencies and police forces. In total, around 80 auditors work on performance audit in the Rekenkamer, auditing €350 billion (Turksema and van der Knaap, 2007). It produces 15 reports a year, recently covering, for example, subsidies and special purpose grants, services for young homeless people, the environmental impact of road transport and jobless people without qualifications. The House of Representatives' Committee on Government Expenditure receives all the Rekenkamer's reports, before they are forwarded to specialist committees. Some time after the publication of the report the Committee collects and passes on questions which members of the House of Representatives have on the audit work. The answers and responses from the audited body are published.

In Sweden, the Riksrevisionen audits the Swedish state, undertaking the financial audit of more than 250 agencies and carrying out about 30 performance audits a year. In 2008–09 this included reports on: controls on cross-compliance in EU farm support; the Government's sale of 8 per cent of the shares in TeliaSonera; cultural grants – effective control and good conditions for innovation; the Swedish for Immigrants programme to provide immigrants with basic language skills; higher education quality; and the quality of private care for the elderly (Riksrevisionen, 2009). Audit findings are presented in a report approved by the responsible Auditor General which is submitted to the government, except where the Swedish parliament oversees the agency. This includes public agencies, foundations, state owned enterprises, entities using public grants and benefits and unemployment benefit funds.

The National Audit Office in the United Kingdom produces around 60 major reports a year, mostly 'value for money' reports. It has statutory powers to examine the economy, efficiency and effectiveness with which central government and a range of other bodies – including universities, further education colleges, hospitals and private contractors working for government – use their funds. Around 270 staff work on VFM audit, many of whom were trained by the office as accountants, but others come from a range of disciplines such as statistics and operational research. Reports by the NAO are used as the basis for hearings of the House of Commons Committee of Public Accounts, which has a wide remit and – in the view of many – a reputation for tough questioning. The Comptroller and Auditor General – head of the NAO – attends the Committee as a permanent

witness, providing, along with the appropriate audit teams, briefing and assistance to the Members of Parliament on the Committee. Reports in the early part of 2010 included topics such as the cost of public service pensions, reorganising central government, Home Office management of major projects, delivering the multi-role tanker aircraft capability, support for business during a recession, regenerating the English regions and the effectiveness of the organic agri-environment scheme.

The United States Government Accountability Office operates under a broad mandate to investigate ‘all matters relating to the receipt, disbursement, and application of public funds’ and ‘to make recommendations looking to greater economy or efficiency in public expenditures’. It is independent of the Federal Executive Branch and reports directly to Congress. In 2007, GAO began work on 554 new performance audit engagements and published 1,038 performance audit products. Around 92 per cent of the performance audits in 2007 were undertaken in response to specific congressional requests or statutory mandates, with the remainder self-initiated. To help provide Congress with information and analysis on a wide variety of questions and enable it to make oversight, policy and funding decisions, GAO produces a wide range of audit products including correspondence, testimonies, briefings and audit reports. Its performance auditors are supported by experts in technical disciplines such as data collection, statistics and law, but also health experts, engineers, environmental specialists and actuaries (GAO, 2008).

Beyond the SAIs

Not all performance audit is undertaken by state audit institutions at national level. Paul Nichol’s examination of audit in Australia (Nichol, 2007) highlights how the states and territories have established varying definitions of performance audit and have different audit laws. Katrien Weets’ chapter in this book focuses on local government performance audit in Rotterdam, one of a number of Dutch cities that have developed their own audit offices. Mark Funkhouser’s chapter also gives examples of performance audit in cities and states in the United States, showing how more and more American cities established performance audit units in the 1980s and 1990s. In the UK, the Audit Commission and District Audit have a long history of value-for-money audit in areas such as local government and the health service, particularly after 1982 legislation, although at the time of writing the government announced the abolition of the Audit Commission and it remains to be seen how local government value for money will progress in the future. Following the introduction of devolved government arrangements in Scotland and Wales, two separate audit

bodies were established – the Wales Audit Office and Audit Scotland. In both cases the form of performance audit developed has differed from the approach followed at the National Audit Office.

WHAT IS THE ROLE OF PERFORMANCE AUDIT?

To date this chapter has emphasised variety and difference, but we can still draw together some common themes which help us to think about the role performance audit plays. Firstly, we can see that it is now widely considered as an essential element of a balanced audit portfolio for an SAI, going beyond, although in some cases feeding off, more traditional audit work. In the UK NAO, around one-third of resources are spent on VFM audit (National Audit Office, 2009). In Sweden the figure is nearly half and in Australia performance audit takes up one-third of the ANAO's resources and 60 per cent of the staff. In many developed audit offices, performance audit holds a substantial, if not the predominant, position, and it is the work which attracts most attention. It is thus an important part of a comprehensive state audit function.

Secondly, we can see the publication of performance audit reports as an accountability mechanism in its own right, offering detailed descriptions or analysis of public sector performance, based on independent access to authoritative documentation and information. Audits generate recommendations on which public officials can, and often do, take action. As Pollitt (2006) has put it:

One type of performance information which does seem to command political attention is that which comes in the shape of performance audits produced by national audit offices. Frequently such attention is semi-mandatory, in the sense that ministers and/or special committees of the legislature are procedurally obliged to respond to such reports. This does not guarantee substantive impact (and certainly not implementation of recommendations) but it does mean that some sort of formal consideration and reply is required.

Performance audits provide visibility to assessments of performance and the increasing willingness of auditors to publicise their work in different ways – via the newspaper, radio and TV, and increasingly on podcasts and even YouTube (the Algemene Rekenkamer's report on elite sports was publicised this way in 2008, as was the UK NAO's work on Successful Commissioning in 2010) – has meant that the work has considerably more profile now than in the 1990s. Audits place in the public domain what Geoff Mulgan has called 'unsettling knowledge' – subjects that governments would prefer in some cases not to have examined or publicised (Mulgan, 2006).

Thirdly, performance audit has increasingly become the basis for legislatures or their scrutiny committees to undertake their work, often providing the majority of the evidence for their enquiries, and for follow up action by government. The relationships between performance auditors and scrutiny committees are many and varied. In the Netherlands, the Algemene Rekenkamer states that the House of Representatives and the government are its principal 'customers' and sees its job 'to provide parliament with useful and relevant information so that it can decide whether a minister's policies are effective. Wherever possible, we match our audits to parliament's wishes and needs.' In the United Kingdom, almost the entirety of the programme of the Committee of Public Accounts is made up of VFM reports (the remainder are mostly reports on qualified financial accounts), which it uses as the basis of questioning senior officials on their organisation's performance. In these settings, as White and Hollingsworth (1999) put it, 'democratic accountability is given real bite: audit allows Parliament to draw aside, at least partly, the veil of ministerial responsibility and to participate in the process by which government controls itself.'

A fourth perspective is that performance audit has enabled SAIs to argue that their work 'adds value' in ways which go beyond concerns with regularity and compliance, which some critics have argued is their very limited perspective (Behn, 2001). Instead, performance audit work has been presented as well considered pointers to how government agencies can improve in keeping with the wider performance agenda referred to earlier. To this end, performance audits have been used, for example, as the basis for conferences and seminars with practitioners in the United Kingdom and the Netherlands, designed to use audit work as a starting point for a constructive and collaborative discussion about performance improvement. Audit work has also generated self-assessment tools for officials and been the basis for considerable informal advice and guidance. Lonsdale and Bechberger examine these themes later. Performance audit has also brought SAIs into close contact with different communities of practice, experts and civil society groups in a way that would not happen if they focused only on the financial accounts. Through this work, SAIs are contributing to wider policy debates.

Finally, in thinking about performance audit, we can see it as one of many forms of policy analysis and evaluation that have grown up in the last 30 years, with its practitioners increasingly tapping into other disciplines in a search for traction. This helps to reinforce the point that performance audit is a hybrid activity, making use of whatever means it can find to generate sufficiently robust evidence for its purposes. The relationship between performance audit and evaluation has been well covered in the audit literature (for example, Pollitt and Summa, 1997, Leeuw,

1996, Pollitt et al., 1999, Mayne, 2006, Bemelmans-Videc, Lonsdale and Perrin, 2007), much of it written by authors with evaluation experience. There is general acceptance that auditors use many of the same methods as evaluators although the way they use them, the institutional settings, the purposes to which the work is often put, as well as the mindset of those undertaking it can often be very different. At the same time, the major advantage that auditors have over evaluators is their greater statutory powers of access and the ready-made, formal processes for ensuring that reports are responded to by government.

THEMES AND STRUCTURE OF THE BOOK

As noted at the start of this chapter we live in a time when concern about public sector performance is at the very centre of political debate. We are also moving from a long period of incremental growth (and occasionally incremental contraction) of the public sector, during which audit offices have honed their skills, to a more turbulent time characterised by wholesale cuts in programmes. Arguably, this could put the audit office and its concerns centre stage of perhaps the most important issues of the coming decade or more. Alternatively, it may render their approaches less useful than in the past, as the issues cease to be about efficiencies, and are more about political decisions around the removal of government from whole sectors. At the same time, the cuts will not pass SAIs by; indeed, the audit offices in Australia and the UK are operating within reduced resources in 2010.

This raises many questions. For example, how should performance auditors respond in their examinations to this new world? Where does the major retrenchment in public spending leave the auditors? If retrenchment is about major cuts in public spending, how can auditors contribute if they do not question policy? It is also clear that, on its own, the proliferation of performance data does not correlate simply with improved performance, but it is less clear how to respond to this. And if the growth of state audit institutions promoting performance audit has not necessarily led to a greater sense of accountability, have their efforts over the last 25 to 30 years been wasted?

As well as being a topical subject, there are other reasons for studying performance audit now. Despite some growth in the last decade or so, several of the main texts are a decade or more old. Yet this period has seen a wealth of case study material which is available to explore aspects of performance audit and, we believe, there is an appetite for a more contemporary analysis based on these studies. At the same time there have been

major developments in the ways in which the work is conducted and considerably more interest shown in many countries in the conclusions and wider contribution of audit work. It has also been argued (RAND, 2009) that ‘in recent decades this task of supporting and delivering performance audit has become technically more demanding.’ In particular:

the architecture of the contemporary state is changing in ways that problematise the role of performance audit and that at least part of the response to this must be to adopt a more sophisticated set of audit practices. In this context, ‘architecture’ refers to the relationships amongst the organisations involved in delivering, communicating and acting upon performance audit (audit bodies, organisations commissioned to support the work of audit bodies, parliament, government, the press, and departments and agencies), the resources they use (money, statutory powers, skills, influence) and power relationships that hold them together.

These changes are considered to be: the problem of agency and identifying who is responsible; the problem of attribution; the problem of measurement; and the problem of ‘whose benefit are we measuring’ and the need to recognise that there may be multiple stakeholders.

Another justification for examining the area afresh is that it is more than a decade since the publication of Power’s book *The Audit Society* (1997), which raised concerns about the impact – sometimes adverse – of audit, and comes after perhaps twenty years of expansion of regulation, evaluation, inspection and scrutiny activity of all kinds. It is, therefore, a good time to be examining afresh the ways in which SAIs operate, how their work is communicated, and with what consequences.

The book is an edited collection, produced by authors who have worked in or around audit offices for many years, and have written about them before. Several are academics or have links with academic institutions. Thus, the book has the perspective of both practitioners conversant with the academic literature on audit, and academics who have been involved in the practice of audit. It is also designed to be of interest to those who are the subject of performance audits – officials in public-sector bodies around the world – and others who use the reports for accountability and performance improvement purposes.

The book is structured into four parts. Accompanying this Introduction in the first part, Jan-Eric Furubo examines further what we mean by performance audit and seeks to identify the essence of the work. His chapter provides a working definition for performance audit for the book, to which we return in the Conclusions chapter. In the second part, we consider how performance audit is conducted in a number of countries. We start by looking at how audit bodies select their studies, a key element of

performance auditors' independence. Vital Put and Rudi Turksema, who both have experience in audit offices, examine the influences on study selection in the Netherlands and Belgium, countries with contrasting approaches to undertaking audits. Vital Put then considers the norms and criteria used in performance audits, against which auditors assess the performance of the bodies they scrutinise. This chapter looks at the norms chosen by audit bodies in the Netherlands and the United Kingdom.

Having selected a topic and decided how to make an assessment, the next stage for performance auditors is to choose the methods with which to gather or analyse the evidence. Jeremy Lonsdale, looking specifically at the United Kingdom, considers how the choice of methods used has developed over the last 20 years at the NAO, focusing in particular on how the context in which the work is done shapes the decisions about what approaches to take. Following this, Alex Scharaschkin looks at how auditors use evidence to structure their arguments. This chapter examines what auditors consider adequate evidence and how they go about marshalling it in ways that will be robust and allow them to withstand challenge from those subject to audit. Linked to this, Justin Keen, an academic in the health field, who spent two years as Research Fellow at the UK National Audit Office in the late 1990s, considers how auditors come to judgements. Performance audit involves the generation and analysis of evidence from a range of different sources. This chapter takes a detailed look at the thought processes and influences on auditors as they reach their judgements, before setting a framework for discussing the strategies that performance auditors use to produce their reports. Finally in this section on the conduct of audit, Peter Wilkins and Richard Boyle consider issues of Standards and Quality. The quality of performance audits is crucial to the standing and credibility of state audit institutions. Over the years, auditors have developed their own standards and created quality assurance arrangements to help maintain and raise standards, but others are now becoming involved in standard-setting processes.

The third part of the book assesses what performance audit is contributing to the improvement of government. This is a contested area. Eddy van Loocke and Vital Put examine the empirical evidence of the impact of performance audits, drawing on 14 empirical studies and concluding that there is evidence that the impact can at times be significant, at others non-existent. Mark Funkhouser – a long-time elected auditor in Kansas City in the United States – follows on in a positive vein and takes on two academic critics of performance audit, Robert Behn and Melvin Dubnick, arguing that auditors have strong grounds for believing that their work has an impact. This chapter shows how, despite their scepticism, there is evidence from the United States that performance audits can and do have real impact on the ways in which government operates. A more sceptical

viewpoint is provided by Frans Leeuw, with a perspective which straddles audit, government, evaluation and academia. He considers the evidence of the perverse effects of performance audit. This chapter – drawing on a wide range of literature – presents the author’s views on problems with performance audit. All three chapters highlight the importance of performance auditors knowing the value of their work, and this is taken further by Katrien Weets, who considers the effectiveness of performance audit at the local government level. Adapting a methodology developed elsewhere, this chapter explores the extent to which performance audits are truly effective, and considers what factors influence their ability to secure impact. It focuses in particular on the work of the city audit office in Rotterdam in the Netherlands.

Continuing on the subject of impact, we end this section with two chapters exploring ways in which performance auditors are seeking to enhance the value of their work – through sharing learning and by being responsive to those they audit. Jeremy Lonsdale and Elena Bechberger look at how performance auditors aim to secure learning, at the same time as playing a role, in accountability processes. In recent years, auditors in many countries have concluded that simply producing audit reports and publicising their findings is not enough to help generate beneficial change. This chapter examines the development of other forms of output from audits, including good practice guides and toolkits designed to assist government officials to do their job better. Peter van der Knaap considers how SAIs can increase their responsiveness to those they audit in an attempt to make their work more effective, looking at how the Algemene Rekenkamer in the Netherlands has been more participatory in its selection of audits, choice of methods and style of audit.

Finally, part four concludes by drawing on the preceding chapters to answer the question – so what? Jeremy Lonsdale, Tom Ling and Peter Wilkins consider the implications of the findings for audit offices and governments, and conclude on the extent to which performance audit is helping to improve the performance of government.

REFERENCES

- AFROSAI (2009), ‘About us’, available at: www.performanceaudit.afrosai-e.org.za/content/about-us (accessed 15 May 2010).
- Algemene Rekenkamer (1989), *Algemene Rekenkamer: Past, Present and Future*, The Hague: Algemene Rekenkamer.
- Algemene Rekenkamer (2009), ‘Performance audits’, available at: www.courtsofaudit.com/english/Organisation/What_do_we_do/Performance_audits (accessed 15 May 2010).

- Algemene Rekenkamer (2010), 'The history of performance auditing at the Netherlands Court of Audit', available at: www.courttofaudit.com/english/Organisation/What_do_we_do/International_act.
- ANAO (2009), The Auditor-General Annual Report 2008–09, available at: www.anao.gov.au (accessed 10 March 2010).
- Astill, S and E Page (2006), 'Aims, instruments and interactions in VFM audit: a cross-national study', Paper for the National Audit Office, London.
- Behn, Robert (2001), *Rethinking Democratic Accountability*, Washington DC: Brookings.
- Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale and Burt Perrin (2007), *Making Accountability Work: Dilemmas for Evaluation and for Accountability*, New Jersey: Transaction Publishers.
- Bouckaert, Gert and John Halligan (2008), *Managing Performance: International Comparisons*, London: Routledge.
- Bowerman, M (1996), 'The rise and fall of VFM auditing', in I Lapsley and F Mitchell (eds), *Accounting and Performance Management: Issues in the Private and Public Sectors*, London: Paul Chapman Publishing, pp. 193–12.
- De Lancer Julnes, P (2006), 'Performance measurement: an effective tool for government accountability? The debate goes on', *Evaluation*, **12** (2), 219–35.
- GAO (2008), 'International peer review of the performance audit practice of the United States Government Accountability Office, peer review report', available at www.gao.gov/about/internationalpeerreviewrpt2008.pdf (accessed 15 May 2010).
- Gronlund, A, F Svardsten and P Ohman (2009), 'Value for money and the rule of law: the (new) performance audit in Sweden', *International Journal of Public Sector Management*, November.
- Gunvaldsen, J and R Karlsen (1999), 'The auditor as evaluator: how to remain an influential force in the political landscape', *Evaluation*, **5** (4), 458–67.
- INTOSAI (2004), 'International organisation of Supreme Audit Institutions ISSAI 3000, implementation guidelines for performance auditing', available at [http://www.issai.org/media\(708,1033\)/ISSAI_3000E.pdf](http://www.issai.org/media(708,1033)/ISSAI_3000E.pdf) (accessed 15 May 2010).
- Lapsley, I and J Lonsdale (2010), 'The audit society: helping to develop or undermine trust in government?', in M Adler (ed.), *Administrative Justice in Context*, Oxford: Hart Publishing.
- Leeuw, F (1996), 'Auditing and evaluation: bridging a gap, worlds to meet?', in C Wisler (ed.), *Evaluation and Auditing: Prospects for Convergence*, New Directions in Evaluation No 71, San Francisco: Jossey-Bass Publishers, pp. 55–61.
- Lonsdale, J (2008), 'Balancing independence and responsiveness: a practitioner perspective on the relationships shaping performance audit', *Evaluation*, **14** (2), 227–48.
- Mayne, J (2006), 'Audit and evaluation in public management: challenges, reforms and different roles', *The Canadian Journal of Program Evaluation*, **21** (1), 11–45.
- Mulgan, Geoff (2006), *Good and Power: The Ideals and Betrayals of Government*, London: Penguin.
- National Audit Office (2005), *State Audit in the European Union*, 3rd edition, London: National Audit Office.
- National Audit Office (2009), *Annual Report*, London: National Audit Office.
- Nichol, Paul (2007), *Audit in a Democracy: The Australian Model of Public Sector Audit and its Application to Emerging Markets*, Aldershot: Ashgate.

- OAG, Canada (2010), 'Performance audits', available at: www.oag-bvg.gc.ca/internet/English/au_fs_e_9365.html (accessed 21 April 2010).
- Pollitt, C (2006), 'Performance information for democracy. The missing link?', *Evaluation*, **12** (1), 38–55.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilkka Summa and Marit Waerness (1999), *Performance or Compliance? Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- Pollitt, C and H Summa (1997), 'Evaluation and the work of Supreme Audit Institutions', in M-C Kessler et al. (eds), *Evaluation des politiques publiques*, Paris: l'Harmattan.
- Power, Michael (1997), *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press.
- RAND Europe (2009), *Performance Audit Handbook*, Cambridge: Rand Europe.
- Rawnsley, A (2009), 'The long party is over for the public sector, whoever wins', in *The Observer*, London, 22 March.
- Riksrevisionen (2009), Annual Report of the Auditors General, Stockholm: Riksrevisionen, available at: www.riksrevisionen.se/upload/1801/annualreport.
- Shand, David and Paul Anand (1996), *Performance Auditing in the Public Sector: Approaches and Issues in OECD Member Countries*, OECD, Paris: Performance Auditing and Public Sector Modernization.
- SNAO (2008), *Audit Plan 2008*, Stockholm: Swedish National Audit Office.
- Talbot, Colin (2010), *Theories of Performance: Organizational and Service Improvement in the Public Domain*, Oxford: Oxford University Press.
- Treasury (2007), *Meeting the Aspirations of the British People, 2007 Pre-budget Report and Comprehensive Spending Review*, London: HM Treasury, Cm 7227, available at: www.official-documents.gov/document/cm72/7227/7227.pdf (accessed 21 April 2010).
- Treasury (2009), 2009 Value for money update, London: HM Treasury, available at www.hm-treasury.gov.uk/d/vfm_2009update_pu787_230409.pdf (accessed 21 April 2010).
- Turksema, R and P van der Knaap (2007), 'Planning and selecting performance audits at the Netherlands Court of Audit', *International Journal of Government Auditing*, July.
- White, Fidelma and Kathryn Hollingsworth (1999), *Audit, Accountability and Government*, Oxford: Oxford University Press.

2. Performance auditing: audit or misnomer?

Jan-Eric Furubo

Performance auditing is thus something of a misnomer. Performance auditing is an evaluative activity of a particular sort whose name happens to include the word 'auditing'. (Barzelay, 1996.)

The importance of performance auditing is explained in textbooks, constitutional discussions and international agreements.¹ It is easy to find powerful advocates such as the World Bank, OECD and the European Union and, as we have just seen, it is now regarded as an important part of the architecture of democratic states. On the other hand, we can find just a handful of books and no journal specifically about performance auditing. We also observe the absence of societies and international meetings for professional discussion about performance audit (although there has been an initiative to found an international centre for performance auditing in the USA in 2009). In contrast, other parts of the democratic structure are subject to seemingly endless discussion.

One possible explanation for this situation is that performance auditing has always been defined by an institutional context. When performance auditing is considered, the discussion centres on what is actually done – or sometimes what is said to be done – in institutions tasked with conducting performance auditing at a certain governmental level (national, state, municipality). Traditionally, though, many of these institutions have been wary of opening themselves up to too much scrutiny of how they operate (Pollitt et al., 1999).

Such a point of departure is tricky. It means that we cannot debate performance auditing in terms other than what a particular kind of institution chooses to call performance auditing. The discussion has, therefore, been very much about Supreme Audit Institutions (SAIs), their actual praxis, and how differences in praxis can be explained. It has also focused on how performance auditing may or may not be related to other forms of intellectual inquiry. There has, in particular, been a tendency to regard performance auditing as a part of the broader landscape of evaluation,

and defined as auditing only because it is conducted by SAIs. It is in a way a circular definition. If we ask a performance-auditing function in a SAI what they do, they will answer 'performance auditing'. And if we ask what performance auditing is, they will answer: 'It is what we do.'

This chapter has a different point of departure. Performance auditing can be, and is, conducted in different institutional settings, a point that is clear, for example, from Katrien Weets' chapter later in this book on audit in Dutch city government. So, the fact that we call something performance auditing means that we imply salient features which can distinguish it from other forms of inquiry. If these features are present we should be able to define it as performance auditing irrespective of which institution conducts the work. And if institutions that define themselves as audit institutions undertake work which has very little to do with these features, the institutional setting is not sufficient reason to call it performance auditing. To identify these features we will consider performance auditing as an 'ideal type' in the Weberian sense. First though, we will discuss briefly why it is important to develop a deeper understanding of what performance audit is all about.

AN UNDERDEVELOPED DISCOURSE

As has already been noted, the volume of what can be described as research and debate about performance auditing is very limited. This is very different compared to, for example, evaluation, or many other aspects of government. This general conclusion is not contradicted even in acknowledging that we do have research about particular aspects of performance auditing (for example Keen, 1999, Morin, 2001, Put, 2005 and Sharma, 2007), and about more general questions in particular countries (for example Ahlbäck, 1999 and Guthie & Parker, 1999).

In addition, much of the literature actually centres on comparisons – differences and similarities – between performance auditing and what many consider is a related discipline, evaluation. One explanation for this is that a lot of the debate has been initiated by players in the field of evaluation and has taken place in journals or other settings with an audience of evaluators. This was obviously the case when New Directions for Evaluation published a special issue *Evaluation and Auditing: Prospects for Convergence* (Wisler, 1996), or, more recently, in John Mayne's insightful essay 'Audit and evaluation in public management: challenges, reforms and different roles' in the *Canadian Journal of Program Evaluation* (Mayne, 2006). The lack of performance audit journals has, in other words, created a situation where the few performance auditors who want to participate

in a wider discussion have often done so through arenas ‘borrowed’ from other professional and intellectual communities, and as a result appear to feel the need to construct their discussion in a way that fits the assumed demands of these journals or conferences (for example Lonsdale, 2008). The lack of a distinctive arena for debate about performance auditing can thus be said to have made performance auditing open for a sort of colonization by the much more dominating evaluation discourse.

It is also the case that many of the authors of contributions on performance audit have had personal experience in audit institutions, as well as more general evaluation environments (universities, governmental institutions, consultancy firms, and so on). This was certainly so for one of the most quoted authors in the field of performance auditing, Eleanor Chelimsky (Oral History Project Team, 2009). The only broad comparative study about performance auditing in different countries was prepared by a team, of which four of the six authors were employed in SAIs, but several had other experiences in evaluation. The leader of the project (Christopher Pollitt) was the first president of the European Evaluation Society (Pollitt et al., 1999). It seems as if this experience of different ‘worlds’ is a driving force behind reflection about similarities or differences related to performance audit.²

In considering the content of the literature on performance auditing, it is possible to distinguish between at least four aspects in the analytical efforts. These are:

1. *Fundamental ideas and concepts*: the role of performance auditing, the underlying rationale for it, as well as consideration of the human and other forces in administrative and political systems.
2. *Differences in institutional pre-conditions*: legal constraints, access to information etc.
3. *The actual praxis*: methods or techniques that are adopted, the objects studied, questions that are asked and norms applied, as well as the degree to which findings delivered by the audit institutions are used.
4. *How this praxis can be explained*: how praxis differs in different jurisdictions and cultures.

From our review it seems as if debate and research have focused primarily on the last three aspects, and probably mostly on aspect 3, with many authors giving us an important understanding of the actual praxis. Schwartz’s discussion (1999) about different strategies by state auditors to cope with the ‘effectiveness dilemma’, for example, is based on an empirical study of the work in six audit institutions. The comparative work by Pollitt et al. covers aspects 2 to 4, and to some extent also the first aspect.

However, few contributions actually deal with what we might call the fundamental ideas and concepts of performance auditing. This does not mean that these questions have been totally ignored; they are to some extent addressed by Leeuw (1996), Barzelay (1996), Mayne (2006) Lonsdale & Bemelmans-Videc (2007), Ling (2007) and elsewhere. But despite this work, we can safely say the discourse about performance auditing is more oriented to praxis and how differences in praxis can be explained, rather than fundamental ideas and concepts. In addition, questions at this level are also many times touched upon *en passant*, before moving to other analytical aspects.

WHY IT IS IMPORTANT TO ADDRESS THE FUNDAMENTAL IDEAS AND CONCEPTS OF PERFORMANCE AUDITING

Even if the reader accepts this picture, it is reasonable to ask why it is a problem, and why it might be important to address the conceptual questions of performance auditing. Is there a risk that such discussion could become a bit excessive? After all, people do conduct performance audits, reports are published and recommendations are acted upon. Why is it essential that there is a reflective literature?

We would argue, however, that developing a discourse in the field of performance auditing, separate from evaluation, is more important today than it was 10 or 20 years ago. The reason for this is to do with the environment in which the work is undertaken. In those countries which developed performance audit in the 1970s and 1980s, performance auditing became an important player in the field of evaluation. In the USA, Netherlands, Canada and Sweden, audit institutions had an important role as producer of evaluations and an important position in the discussion about evaluation (Mayne et al., 1992, Gray et al., 1993, OECD, 1996). The role of the audit institutions as pioneers and entrepreneurs for evaluation at that time is emphasized in several studies (Furubo et al., 2002).³ Today the situation is different. We have a stream of evaluative information produced by many different institutions (Rist & Stame, 2007). The performance audit departments within SAIs produce a very limited share of this evaluative information in those countries with a long performance auditing tradition. Think tanks, research centres and consultancy firms all offer reviews of public policy initiatives. And in countries that have developed a performance auditing praxis more recently, it is unlikely that performance auditing will have as dominating a role as it had in a handful of countries a couple of decades ago.

So, today, when performance audit is just one of many producers of evaluative information, very often asking the same questions as other producers, but protected and bolstered by specific constitutional arrangements, it is important to establish the fundamental ideas and concepts of performance auditing so we can be clear about what it has to offer that is unique.

DIFFERENT INTERPRETATIONS OF PERFORMANCE AUDITING

A discussion of these ideas and concepts is immediately confronted with the fact that it is not easy to capture the meaning of performance auditing. It might be assumed that a natural point of departure is that it is a specific form of auditing; after all it is called 'auditing'. However, even here we can easily find contradictory opinions. Barzelay, for example, states that 'Performance auditing is thus something of a misnomer. Performance auditing is an evaluative activity of a particular sort whose name happens to include the word "auditing"' (1996, 19). And, at a more anecdotal level, many practitioners who have come to performance auditing from backgrounds other than traditional/financial auditing may well assure any listener that even if they do something called performance auditing, they are certainly not auditors. They are evaluators, analysts or preparers of briefing material, and so on but certainly not auditors.

Given this interpretation, the term performance auditing can be regarded as a form of 'conceptual stretching' (Sartori, 1970, 1033). In this way, when audit offices extended the traditional territory of financial auditing, the expansion was facilitated and legitimated by naming it auditing. In this context it is perhaps even more interesting to observe the ambiguity in the descriptions of performance auditing within the work of the International Organization of Supreme Audit Institutions (INTOSAI). Its Auditing Standards state that 'Performance auditing is concerned with the audit of economy, efficiency and effectiveness'. Its implementation guidelines for performance auditing (2004), however, talk of the 'independent examination of the efficiency and effectiveness of governmental undertakings . . .' The guidelines also state that 'Performance auditing does not have its roots in the form of auditing common to the private sector. Its roots lie in the need for independent, wide-ranging analyses of the economy, efficiency, and effectiveness of governmental programs and agencies made on a non-recurring basis' (INTOSAI, 2004, 12).

We must assume that the fact that the guidelines avoid defining the *activities* conducted by performance auditors as auditing, and instead

use words such as ‘examination’ and ‘analysis’ is not something that has happened by chance. The different perspectives in the guidelines issued by INTOSAI are made very explicit in the following excerpt:

Auditing is normally associated with accountability, but in performance auditing this is not always the case. Auditing [for] accountability can be described as judging how well those responsible at different levels have reached relevant goals and met other requirements for which they are fully accountable (factors outside the control of the auditees are not expected to influence the outcome). An alternative approach is to focus on understanding and explaining the actual observations that have been made during the audit. Instead of trying to find out who is at fault, it is possible to analyse the factors behind the problems uncovered and to discuss what may be done about these problems. This approach reflects the idea that the overall aim of performance auditing is to promote economy, efficiency and effectiveness. The two approaches represent different ideas to performance auditing; one in which accountability (as it is in compliance and financial auditing) is at the centre of attention of the audit, while the other – which put emphasis on economy, efficiency and effectiveness – primarily concerns itself with the subject matter of the audit cases of problems observed. (INTOSAI, 2004, 28)

It is also worth noting that until recently the NAO in the United Kingdom talked of ‘value for money examinations’, rather than audits, at a time when it was introducing a wider range of methods and perspectives into the work, but recently has taken to using the word ‘audit’ again to emphasize the proximity of the work to its other main activity – the audit of financial statements. This reflects a desire to bring the two activities closer together again after 20 years of growing apart.

We will return to these perspectives, but what has been demonstrated is that the assumption that performance auditing is auditing is certainly not self-evident. In the following sections we will therefore try to give a deeper understanding of the relationship between *auditing* and *performance auditing*.

PERFORMANCE AUDITING CONSTRUCTED

The first step in constructing the essential elements of performance audit is to describe what auditing is. As a second step, we will consider if we need to take account of other operations beyond those associated with traditional (or financial) audit. By doing this we can consider whether performance auditing should be regarded in the same way as we regard other forms of auditing. Thereafter, we will discuss how performance auditing can be defined.

Step 1: Auditing as Ideal Type

In order to answer the question of what auditing is, we use an instrument of analysis which we associate with Max Weber's name, namely the 'Ideal type'. It is important to emphasize that Weber's Ideal type is a theoretical construction aimed at creating conceptual clarity from a particular point of view. It is not a concrete representation of the real world or a normative model (Brante et al., 2001, 125). Weber himself stated in 1904 that 'An ideal type is formed by the one-sided accentuation of one or more points of view and by the synthesis of a great many diffuse, discrete, more or less present and occasionally absent concrete individual phenomena, which are arranged according to those onesidedly emphasized viewpoints into a unified analytical construct (*Gedankenbild*). In its conceptual purity, this mental construct (*Gedankenbild*) cannot be found empirically anywhere in reality' (Weber, 1949, 90). Weber also emphasized 'that the idea of an ethical *imperative*, of a "model" of what "ought" to exist is to be carefully distinguished from the analytical construct, which is "ideal" in the strictly logical sense of the term' (Weber, 1949, 91).

The difficulties of defining a phenomenon based upon how a certain term is used as a label for different praxis are illustrated in Power's discussion (1997) about auditing. He states that 'The ambiguity of auditing is not a *methodological* problem but a *substantive* fact' (italics in original, 1997, 6). So, it seems it is not straightforward to assert what audit actually is if we try to understand auditing from its concrete realizations in different forms of the work.

To figure out what audit means in terms of an ideal type we must go back to what we can regard as the fundamentals of auditing. In his *Philosophy and Principles of Auditing*, Flint proposes a number of postulates on which a theory of auditing can be constructed. He expresses his first postulate in the following way: 'The primary condition for an audit is that there is a relationship of accountability or a situation of public accountability'. He elaborates on this by sub-dividing it into several parts and into different forms of accountability. The first is that there is a 'relationship of accountability between two or more parties in the sense that there is a duty of acceptable conduct owed by one party to the other party or parties'. The simplest form of this postulate is a situation with two specifically identified parties (Flint, 1988, 23).

Flint points out that the essential element in all these situations is 'that the actions of the first party . . . are subject to scrutiny by another group or groups who are in the position to invoke or impose some sanction or penalty if the performance falls short of what can be expected by reference to the criteria that have been established. The first parties required to give

an account, provide information or provide access to information as a basis of decision by the group or the groups to whom they are accountable' (Flint, 1988, 24). This relation can exist on several levels and can be seen as a chain of accountability relations. In his book, *Rethinking Democratic Accountability*, Behn talks about accountability holders or holdees, and it is possible to be an accountability holder in one relationship and an accountability holdee in relation to a superior level.

However, we can also talk about accountability in relationships other than that between an agent and a principal, defined as somebody who can 'invoke or impose some sanction'. It can also be about the need 'by some party to establish the credibility and reliability of information for which they are responsible which is expected to be used and relied on by a specified group or groups of which the members may not be constant or individually identifiable' (Flint, 1988, 22). Some of these groups are easily identifiable such as prospective investors, employees, parliamentarians, and so on. The point is that we are talking about entities that have a responsibility to report information and groups which have the right – by tradition or law – to request information from another party and also to demand that this information is verified in some way.

We can note that the first postulate does not say anything about the actual need for auditors or audit functions. Flint's second postulate rectifies this, stating: 'The subject matter of accountability is too remote, too complex and/or of too great significance for the discharge of the duty to be demonstrated without the process of audit' (Flint, 1988, 26). Even if we can imagine a relationship where the accountability holder himself can scrutinize the accountability holdees, it is probably easier to imagine such situations in which we need a third party – the auditor – to conduct this scrutiny. We can also note that the independence of this third party is seen as a pre-requisite by Flint. This seems necessary given the first two postulates. The function of the auditor in the accountability process is based on independence and freedom from investigatory and reporting constraints.

Step 2: To Include Performance Auditing in Auditing

Weber's ideal type and its one-sided accentuation of one or more points of view means that it is constructed from a certain perspective. It means that different perspectives regarding the same phenomenon can lead to the construction of different ideal types. So far, the points of view or perspectives have been uncontroversial. We will – more or less – find the same view in most textbooks in which auditing in general is dealt with. Auditing is about a relationship of accountability and the need for a third

– independent – party. Few will dispute the need for auditing, described in that way, when it comes to scrutinizing financial statements which present the pattern of financial transactions and the overall financial position of the entity and to ensure that public money is not spent for private purposes. Most of us, even if we basically have a very positive view of human nature, accept that it is always a risk that people will embezzle or demonstrate other forms of dishonest behaviour.

It is when we talk about performance in more general terms that the question of whether we actually need auditing becomes a bit confusing and complex. We can ask, like INTOSAI itself has, if we really need to scrutinize the performance of agents in the same way it is intended they be scrutinized in financial auditing. We can also ask if it is not sufficient to tell decision makers and managers how they can improve things. We can further ask if it is not more important to discuss what may be done about the problems which are discovered rather than blame individuals or institutions. Such questions reflect underlying assumptions about individuals and administrative systems. On the one hand we might assume there is a risk that people will embezzle or take bribes. On the other hand we might assume that people in all other respects are willing to loyally perform their duties in a way that reflects the intentions of their principals. We might also assume that they are willing to learn and improve in order to pursue efficiency and effectiveness in the public interest.

These contradictory assumptions about bureaucratic systems and individuals within these systems are of fundamental importance to how we regard performance auditing. In order to refine possible points of view we will make use of a nearly seventy-year-old debate between Carl J. Friedrich and Herman Finer. In the Friedrich and Finer debate, two different notions about public administration clashed. Using colloquial terms the clash was between a ‘good guys’ and a ‘bad guys’ notion. In this debate Friedrich and Finer agreed about the growing discretionary power of the administrative systems. Friedrich’s recipe for reconciling this development with democracy included two ingredients; technical knowledge and political responsibility. Technical knowledge means, according to Friedrich, that ‘administrative officials seeking to apply scientific “standards” have to account for their action in terms of a somewhat rationalized and previously established set of hypotheses. Any deviation from these hypotheses will be subjected to thorough scrutiny by their colleagues in what is known as “the fellowship of science”.’ Political responsibility meant that the administrator should not only take into account existing knowledge but also ‘existing preferences in the community, and more particularly the prevailing majority’ (Friedrich, 1965, 169). Friedrich assumed that it was possible to create a sort of internal corrective mechanism and

internalize different values, which means that it is possible to assume that bureaucratic systems – and individuals – will act as ‘good guys’.

Finer advocates the opposite position and takes a darker, more ‘bad guys’ oriented perspective. The discretionary power has increased, but the response to this development is to increase the elected representatives’ possibilities to ‘determine the course of action of the public servants to the most minute degree that is technically feasible’ (Finer, 1965, 77). Finer’s position is based on the perception that bureaucratic structures have to be controlled. The reason why we have to regard administration with such mistrust is that ‘the political and administrative history of all ages’ has demonstrated ‘without the shadow of doubt’ that in all forms of government, including the benevolent ones ‘sooner or later there is an abuse of power when external punitive controls are lacking’. This abuse of power has according to Finer shown itself ‘roughly in three ways’ (Finer, 1965, 179 f).

- Government and officials have been guilty of nonfeasance, that is to say, they have not done what law and custom required them to do owing to laziness, ignorance, or want of care for their charges, or corrupt influence.
- Again, there may be malfeasance, where a duty is carried out, but is carried out with damaging consequences, because of ignorance, negligence, and technical incompetence.
- Third, there is what may be called overfeasance, where a duty is undertaken beyond what law and custom oblige or empower; overfeasance may result from dictatorial temper, the vanity and ambition of the jack in office, or genuine, sincere, public-spirited zeal. Finer emphasizes the importance of this source of abuse and states that ‘we in public administration must beware of the too good man as well as the too bad’ (Finer, 1965, 180).

Finer’s perception of human nature is obviously dark. However, in a debate about auditing it is not important to decide which of these positions, the one held by Friedrich or the one held by Finer, is the most valid description of human nature. The reason is simple. Most of us will probably accept one of the following statements.

- 1) Friedrich’s description is basically correct. In spite of this, we need external control systems, either because other moral mechanisms will not work perfectly in all situations and/or because we will sometimes be confronted with exceptional situations or exceptionally (bad) individuals.

- 2) Finer held a basically more realistic view. In spite of this, we can also at the same time observe a fundamental altruism and willingness to fulfil the political intentions using the best technical knowledge. This means that many individual bureaucrats as well as bureaucratic systems will use new knowledge for improvement and 'social betterment'.

This line of reasoning means that irrespective of which position is most 'right' we will, in practice, act as if both were correct. It means that the knowledge produced about governmental activities can have different purposes. A lot of what is produced, and a big proportion of the stream of evaluative information, can be placed at the 'good-guy-point' on a fictitious scale. But it also means that we need information resulting from scrutiny of the behaviour of administrators and administrative organizations based on the assumption that 'there is an abuse of power when external punitive controls are lacking'.

However, accepting the proposition that flaws in human nature which have made it important to have auditors to discover, and hopefully prevent, dishonest behaviour in pecuniary matters, leads to accepting the necessity of widening the scope of auditing. By this argument, economy, efficiency and effectiveness will also be under threat due to fundamental flaws in bureaucratic systems and human nature, and we cannot trust bureaucrats and organizations to be interested in describing their own shortcomings in order to improve. This insight, offered by many different scholars, is summarized by Schwartz who states that 'Administrators' interest in organizational stability, budget maximization and the promotion of a favourable image contribute to a general desire to refrain from conducting evaluations which might show agency programmes in a bad light' (Schwartz, 1999, 511).

So, irrespective of how we describe the actual praxis of the performance auditing conducted by existing auditing institutions, it is possible to argue that we need auditors to cover many different aspects of the performance of governmental institutions. Such a standpoint implies that performance auditing is auditing, and that the *raison d'être* of performance auditing is the same as in other forms of auditing, namely to scrutinize the fulfilment of different tasks in a context of accountability. In this process the centre of attention can certainly be the three Es (economy, efficiency and effectiveness) often mentioned in relation to performance auditing, but also many other aspects of organizational performance.

What we have highlighted is that auditing is based on certain assumptions about human nature and the nature of bureaucratic structures. And, in relation to performance auditing, it is important to point out that these

assumptions can differ from the assumptions on which other forms of evaluative praxis are based. In other words, it is possible to distinguish between different theories which can explain the impact of different forms of information or knowledge.

The assumed mechanism for many evaluation theorists seems to be that professionals and decision-makers are willing to use knowledge, and that the motives of public administrators are the common interest, rather than the self-interest. The belief is that people who are given knowledge will make better decisions and will act more effectively than people without this knowledge. Mark, Henry and Julnes point out that 'The *raison d'être* of evaluation . . . is to contribute indirectly to social betterment by providing assisted sense-making to democratic institutions . . .' (Mark et al., 2000, 7). The 'good guy' notion can also be regarded as central in many of the ideas about public administration which have been widely embraced in the last two or three decades with labels such as Management by Objectives, Performance Budgeting, Results Based Management and, of course, New Public Management (NPM) from the 1980s onwards.

In contrast, we have the much darker notion of human nature which here has been represented by *Finer*. Even from this perspective, the studies of different programmes or activities of different bodies can contribute to improvement, but it is due to the existence of somebody who seeks to scrutinize and control. Control supports desirable values and good behaviours in society, which in turn leads to better decisions and better use of resources, and thereby has an impact on the level of trust in society. It can be seen as a paradox, but because audit is based on distrust it will contribute to trust. Hence the argument runs that because I fundamentally distrust government it will be possible to trust government only when I know that government is controlled. This is not a very original idea. Indeed, it is the basis for the construction of many constitutions.

When we broaden the scope of auditing to cover aspects of performance other than those covered by traditional/financial auditing we do so because we think that the 'bad guy notion' is at least partly true. Such a position is also the rationale for the independence of the audit institutions. If everybody is a good guy, independence does not seem so important because everybody can be trusted to carry out their roles without oversight. Because we believe this is not the case, in order to make it possible to trust government we need institutions which scrutinize governmental activities from an independent position. In other words, a basic precondition for control and accountability is independence. Independence therefore has a much more fundamental importance for performance auditing than for evaluation, where evaluation is designed to improve

decisions or praxis. If you are in the business of producing information for a more general debate or for the improvement of individual interventions, independence can be regarded much more conditionally. When it comes to audit institutions, independence is not a quality which *sometimes* can be attributed to individual reports, but is a part of the definition of what audit is always about. Without independence we cannot talk about auditing.

This insight is crucial. Independence can also be attached to other forms of evaluative activities. It can be argued that a study aimed at improving the actual praxis of administrators and different programmes *can* be more trustworthy and thereby more useful if it is produced in a more independent manner. However, such a point of view is conditional. The reason for this is that we can also argue for the opposite. It can easily be argued that dependence, instead of independence, is important in order to enhance use and improvement. Michael Patton, perhaps more than anyone else the person associated with Utilization-Focused Evaluation, argues that ‘. . . utilization-focused evaluation should be judged by its utility and actual use; therefore evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything, from beginning to end, will affect use’ (Patton, 2004, 277). The idea that evaluation should be done independently from the persons or institutions which are studied is certainly some way distant from Patton’s thinking. Independence can, in other words, be discussed in instrumental terms if the rationale for the evaluative efforts is learning or improvement. In contrast, a fundamental part of the characterization of auditing is that it is *always* independent.

To use the word ‘always’ in relation to independence in such a manner can invite objections. It is easy to point out that, after all, nobody is independent. We certainly admit that the actual meaning of independence can be elaborated on. However, what we mean by independence is a great enough degree of freedom from investigatory and reporting constraints that the assessments can be used in the context of accountability. This includes the auditor having the right to choose what to audit, deciding about methods and also how to disseminate findings. But it also has to include a certain capacity to carry out different tasks but also freedom from the fear of punishment in different forms if the auditor reveals truths which are uncomfortable.

Step 3: A Definition of Performance Auditing

Based on the postulates we have summarized and what we consider to be the fundamental conditions for information intended for use in a context of accountability, a definition could be:

Performance auditing is an evaluative activity which produces assessments regarding performance or information about performance, of such a reliable degree and with such a freedom from investigatory and reporting constraints, that they can be used in the realm of accountability.

The definition emphasizes that performance auditing is an evaluative activity. We need therefore to consider what are the key characteristics of such work so that we can see how they apply to performance audit. Following Scriven we can talk about evaluation as the process of determining the merit, worth, or value of something, or the product of that process (Scriven, 1991, 139). An evaluative activity is therefore concerned – in one way or other – with merit, worth or value. It means that evaluation and evaluative information can be defined very broadly.⁴ In many countries, evaluation is regarded as a retrospective analysis, but there are other interpretations, as noted by Chelimsky, who has commented that ‘we now evaluate both *ex ante* and *ex post*’ (Chelimsky, 1995, 4). However, given the context of accountability, ‘evaluative activity’ in our definition focuses on what has been done, or is about existing (or completed) interventions. The definition can therefore be said to imply that we are talking about assessments *ex post*, although this can also encompass existing preconditions for future success.

Most evaluators acknowledge that we have to draw boundaries between evaluation and other forms of intellectual inquiry. For example, descriptive statistics as such cannot be labelled as evaluation. Statistics, which in many countries have been produced for more than 200 years, can be an important part of evaluation and often have fundamental importance for how we conceive society and how we define social problems. But statistics about housing standards, or the number of students in different education systems and so on, cannot be regarded as evaluation. Similarly, a statement about a social problem does not become an evaluation merely because it is linked to an assertion that a policy is ‘stupid’ or ‘brilliantly conceived’. In other words, evaluation has to be a carefully conducted assessment of the merit, worth, and value of the evaluand (Vedung, 1997, 3), rather than an ill-considered statement of a view. When we state that something is an evaluation, we imply some sort of quality in the evaluative product and in the evaluative process. However, such a statement does not mean that evaluation has to be conducted according to a set of specific criteria which reflect certain epistemological perspectives.

An overly vague use of the term, however, makes evaluation more or less synonymous with all forms of knowledge production emanating from academic institutions, consultancy firms and others. While the activities undertaken by the latter groups certainly can be described as systematic

and careful endeavours, there is an additional salient feature of evaluation, namely that it is an investigative activity conducted in relation to something which reflects a particular kind of purpose and intention.

So, our definition of performance auditing as an evaluative activity means that it is one which:

- 1) produces assessments that involve explanations and judgements, not just descriptions;
- 2) is conducted carefully; and
- 3) is related to some form of existing or finished intervention or, to put it more broadly, to something which has been done or is being done.

The word ‘assessments’ can also be seen very broadly. It does not say anything about the actual procedures or how these assessments should be communicated. It can be in ad hoc individual reports, or be communicated in the form of an ongoing stream of information yearly or daily – at least in theory – regarding some predetermined variables.

A fundamental part of the definition is also that these assessments can be used in the realm of accountability. It means that the assessments need to have characteristics which make such a use possible. The assessments must be clear about who is accountable and about the extent to which something which has been described as satisfactory or unsatisfactory can be attributed to the conduct (or lack of such) of certain individuals or organizational bodies. Part of the definition is that accountability can be ascribed to somebody (person or institution), who is committed to answer. There have to be accountability holdees, who can be rewarded or punished on the basis of judgements about their performance. As a result, the information which is produced has to satisfy high demands on reliability and also exactness. Information designed for processes of accountability has to be considered to be the truth and – in one way or other – built around evidence. We will return to the implications of this shortly.

This perspective of accountability also means that the part of the definition concerning independence or ‘freedom from investigatory and reporting constraints’ needs to be discussed further. It is obvious that Flint, from whom we have borrowed the phrase, and others, think primarily about constraints given by the agent. However, in a governmental context, and especially in a parliamentary system, something has to be added. In a parliamentary system the actual relationship between parliament and government can be very distant from the model of a superior principal and a subordinated agent. In real life this relation can even – at least partly – be reversed. In a parliamentarian system the relation between the legislative body and the executive body will therefore risk

being, in a manner of speaking, not a genuine principal–agent relation. The implication of this is that the independence of audit institutions in parliamentary systems cannot only be discussed in terms of independence of the executive system.

The important thing is that information can be used in order to hold individuals or organizations accountable about a programme or activity. The definition also implies that performance auditing can establish the reliability and credibility of information regarding different aspects of performance. Such information can be produced by the body responsible for the implementation of a programme or an activity, but it can also be produced by a separate institution.

The definition does not say anything about the criteria or focus (such as the 3 Es, economy, efficiency and effectiveness) which have to be adopted in order to value the performance. Because performance is a very broad concept, as well as the three 3 Es it can also be about the production of a certain output, the societal effects of such an output, or the quality of the outputs from a variety of perspectives. Equally, it is important to add that the words in the definition ‘used in the realm of accountability’ do not mean that performance auditing assessments always have to allocate blame. If we imagine that a performance audit discovers problems in the implementation of a programme such that the intentions of the programme are not fulfilled, it can be asked to what extent these problems or failure can be explained by the conduct of certain bodies or individuals. And if the problems can be attributed to the conduct of a certain party it has also to be asked to what extent it was possible for this party to make other choices than it actually did.

In conclusion, the definition of performance audit we have set out and debated is not institutional. It is possible to imagine an inspectorate or other oversight body undertaking work to this definition as much as an SAI. We can also imagine that a parliament or a government could give a broad assignment to a consultant firm which also satisfies the definition. Of course, the question of independence and freedom from constraints is a question about degree and at some point the boundaries will narrow the possibilities of operating in a manner in which it is no longer possible to talk of ‘freedom from investigatory and reporting constraint’.

CONSEQUENCES AND CHALLENGES

Our definition of performance auditing has an impact on what information performance auditing actually can produce and how it can be used. We will point out three important problems and consequences.

An Ambiguous Relationship to Politics

Performance auditing has an intimate and complex relationship to politics. The INTOSAI implementation guidelines state that ‘Performance auditing is based on decisions made or goals established by the legislature’ (INTOSAI, 2004, 11). In other words, INTOSAI emphasizes an assumed borderline between policy and implementation. The idea of this borderline can also be observed by examining what SAIs say about themselves. The Algemene Rekenkamer in the Netherlands, for example, emphasizes that ‘we investigate whether central government revenue and expenditures are received and spent correctly and whether central government policy is implemented as intended. . . . We express an opinion on government policy that has already been adopted; we do not express political opinions’ (Algemene Rekenkamer, 2010). The Bundesrechnungshof in Germany states that ‘Performance auditing looks into whether the optimum ratio between the objective pursued and the resources utilised has been sought and obtained. It covers the effectiveness and appropriateness of government operations and transactions including target achievement (effectiveness or programme results audit). It also covers the examination whether the input of resources was kept to the minimum necessary to achieve the preset objectives (principle of economy)’ (Bundesrechnungshof, 2010).

The same picture emerges from the statements of many other Supreme Audit Institutions (EUROSAI, 2005, National Audit Office, 2005), and corresponds with that given by Pollitt et al., who note that: ‘For some SAIs, particularly the NAO, questioning policy is a definite no-go area; for others the line is not so clear’ (Pollitt et al., 1999, 26). This position does not necessarily stand still. Pollitt et al. pointed out in 1999 that in Finland and Sweden ‘the watershed between policy objectives (as a no-go area) and the implementation . . . is not as salient an issue as with the NAO’ (Pollitt et al., 1999, 27). However, the situation in Sweden has since changed; it is now emphasized that the point of departure for a performance audit is the Parliament’s intentions (Riksrevisionen, 2008, 10).

We can thus see that the borderline between politics and implementation is a vivid issue for SAIs, and, in the experience of many performance auditors, a tricky problem in practice. Pollitt et al. state:

In theory, the line between questioning policy objective and only asking questions related to the ‘three Es’ in the implementation of policies may sound a clear cut and unambiguous way of defining the mandate for performance audits. In practice, however, if a performance audit project is extended into analytical questions on *why* the observed effectiveness of an audited programme is unsatisfactory, it may not be easy to keep away from explanations

that raise questions of relevance or realism in the specification of objectives. (Pollitt et al., 1999, 27.)

Part of this thinking stresses that it is through administrative systems that the political level can fulfil its promises to the electorate and that it is possible to judge interventions and to develop a perspective that is somehow value-free. The emphasis on efficiency 'implies that the policy implementation can, indeed be separated from policy decisions. If there exists a one best way (and thus a most efficient way) to carry out any policy decision, and if there exists an organizational apparatus for deploying all of these one best ways, government can, indeed, separate administration from policy' (Behn, 2001, 41). However, Lane suggests that the philosophical argument against the politics/administration separation is extremely weak, as it is not possible to uphold a separation between a number of entities (Lane, 2000, 28): ends versus means; values versus facts; preferences versus instruments; norms versus existence; and directives versus reality.

It is worth noting that the diffusion of performance auditing has taken place during the same period as these forms of dichotomies have been spread and as we have noted, the self-perceptions of the SAIs seem to be based on such dichotomies when they state that they do not express opinions about the merits of different policies and restrict themselves to economy, efficiency and effectiveness. The problem the SAIs face is that a lot speak against this separation of politics and administration. The political often centres around *how* to reach different goals, the instrument, instead of ends or fundamental objectives. For example, few oppose the goal of reducing crime, improving the knowledge of children in maths or reducing the transmission of sexual diseases, but we can have very different ideas about how to achieve these goals.

This is certainly a challenge for performance auditing. To argue that political decisions are a no-go area for performance auditing means that it sometimes will be very limited in scope and will be reduced to the scrutiny of technical implementations. However, to argue for the opposite raises issues of objectivity and suggests the auditor has a political role. What is clear is that different rationales for performance auditing have consequences for how this problem is handled.

Accountability and Learning

Lonsdale and Bemelmans-Videc (2007) differentiate between auditing and other forms of evaluative information in relation to how they 'stand with regard to accountability and learning'. They based their discussion on a number of criteria. Firstly, work that is usually externally imposed,

with the subjects of the work having little or no degree of influence over whether it takes place or not, is more likely to have a strong accountability element. Organizations commissioning evaluations of their own operations are clearly more in control than those subject to statutory audits. Secondly, work where the subjects cannot negotiate the timing of the work (*when* it will take place) will similarly have a strong accountability feel. For example, financial audits usually take place, come what may, at a set time – after the year-end on an annual cycle. Thirdly, the extent to which the subject can negotiate or influence the approach will also be important (*how*) (Lonsdale & Bemelmans-Videc, 2007, 12).

Based on these criteria they express the view that the internal evaluation commissioned for learning purposes, and which is not published, can be placed very high on the learning axis but very low on the accountability axis. More or less the opposite can be said about financial auditing. They place performance auditing (slightly) below financial auditing on the accountability axis and slightly higher on the learning axis. The authors explain the fact that they give performance auditing a value on the learning axis as well as the accountability axis indicates that ‘Although accountability is the main purpose – with the report likely to be the subject of further scrutiny (parliamentary committee, press, and so on) – the report will include recommendations for improvements based on demonstration of weaknesses, and may highlight areas for learning’. A similar discussion is conducted in relation to performance auditing which is ‘recommending improvements and changes’.

It is important to emphasize that Lonsdale and Bemelmans-Videc regard accountability as the main purpose of performance audit, with other effects as a sort of (positive) side-effect. In other words performance auditing does not exist because it is regarded as a learning tool, and the authors argue that other forms of investigative or evaluative activity are more efficient when it comes to learning. The learning effect of studies commissioned and conducted for internal purposes can be higher than of studies conducted by independent institutions. So, to continue the same line of reasoning, if the rationale of performance auditing was learning it could certainly be questioned if audit institutions should be independent!

Reliability versus Outcome/Impact

Given that our definition states that audit is something which takes place in the realm of accountability it is possible to argue that the findings which are delivered by performance auditors should be very reliable. Problems with reliability exist on many levels but are probably more difficult to handle when it comes to studies of the societal effects of certain

interventions. In several audit institutions there has been a tendency to emphasize the importance of such studies, although as Pollitt et al. (1999) and Schwartz (1999) have noted there appears to be a discrepancy between intentions and actual practice. Personal discussions with representatives from performance auditing institutions indicate the same thing.

A reason for this discrepancy is probably the tension between such studies and the demand for reliability. In discussing possible connections between a certain intervention and societal developments which the intervention is designed to secure, we have to imagine different situations. Taking a programme aimed at reducing traffic accidents or improving the health of schoolchildren, a general effect-oriented question is: Has this programme actually reduced traffic accidents or has it improved health?

When we try to establish the connection between what the government has done and the effects in society we deal with three questions.

- (1) What has taken place? In other words: Has the health situation improved or have the traffic accidents decreased?
- (2) To what extent have the actual developments been caused by the interventions?
- (3) Can these observations be used in future decision making? In other words: Can this retrospective information be used in the construction of future interventions?

This chapter is not concerned with the first question, which is a matter of statistics. It is when we have to tackle the following two questions that we find ourselves in a situation which not only involves practical issues but also a fundamental theoretical problem. As the Swedish economist and Nobel Prize Laureate Gunnar Myrdal put it more than 70 years ago, an inquiry into the effects of a regulation 'must relate to the difference between the economic events which should be triggered by the employment of a different regulation. This other regulation (the norm of comparison) must therefore also be specified. In any other sense than as a difference between two sets of events initiated by alternative measures of financial regulation, all talk of effects will be devoid of definite content' (Myrdal, 1934). Recent discussions about evidence in evaluation are very much about these problems (Donaldson et al., 2009). We have to accept that our ability to isolate the effects of a certain intervention from all other factors which influence the actual development is limited.

An acceptance of these problems is not the final blow against the craft of evaluation, but of course it would have been if evaluation was only about establishing the causal links between the programme or the intervention and the actual development. But the actual practice of evaluation also

demonstrates that evaluation also can provide us with answers to many other questions. Evaluation can help us to ask critical questions about the underlying theory for a particular intervention, and help us understand which assumptions the interventions is based on. However, we are in a different situation if we are in the business of accountability. The uncertainties, which many times can be a focus in themselves if we want to create deliberation and a deeper understanding of the social mechanisms, are a problem in the context of accountability. They are a fundamental weakness given the role of the auditor, as an independent third party, to give the accountability holder answers about performance or correctness of information which can be used in a process of accountability.

So a consequence of our definition is that performance auditing has difficulties in making reliable statement is about the impact of complex interventions in complex circumstances, and this probably explains why SAIs have difficulty in handling the third E (effectiveness), in spite of their rhetorical ambitions. Mayne has also taken this discussion in a normative direction when he argues that performance auditing should avoid effectiveness (Mayne, 2006).

CONCLUSIONS

This chapter has identified a form of 'conceptual stretching', although not the stretching of the concept as such, but rather letting the concept encompass new areas. We have argued that performance auditing is auditing and that it maintains the salient features of audit. The difference is that it scrutinizes the fulfilment of tasks wider than traditional auditing; arguing that it is the financial audit foundation that has helped audit institutions to accommodate and legitimate their conduct of investigative activities beyond those associated with traditional financial auditing.

The rationale for audit is that the scrutiny of bureaucratic systems or bureaucrats is needed irrespective of whether we are concerned about the risk of embezzlement and forgery, or the risk of how well the administrators, more generally, handle their many duties. In addition, if we also suspect that organizations tend to suppress unfavourable information about their performance and exaggerate favourable information, it is important that it is possible to hold office-holders accountable for a broader territory than is covered by traditional/financial audit. This perspective certainly paints a dark picture of human nature and bureaucratic systems; the fundamental point of departure being that we have to mistrust both individuals within the systems, as well as organizations they work within. On the other hand, a society which is permeated by distrust

will run into fundamental difficulties. We therefore need mechanisms and institutions which can create trust. Such a fundamental mechanism is effective accountability and control, including the existence of procedures which create confidence that public officials and organizations are held accountable. Thus, we have the paradox: because we distrust the men and women tasked with implementing the will of the people, we need to scrutinize how they fulfil different tasks in order to trust them.

So, performance auditing is based on the same rationale as auditing in general and maintains many of the same features. It facilitates our ability to hold office-holders and others accountable for their responsibilities and also means that independence is part of performance auditing in an unconditional or non-negotiable way. We have further emphasized that performance auditing can be characterized as follows:

- it produces assessments that involve explanations and not only descriptions of a certain development
- it has to be related to some form of existing or completed intervention or activity
- it produces assessments which involve judgements of individuals or organizations, not only of a program or an activity and
- it is carefully conducted.

Thus our definition of performance auditing is:

Performance auditing is an evaluative activity which produces assessments regarding performance, or information about performance, of such a reliable degree and with such a freedom from investigatory and reporting constraints, that they can be used in the realm of accountability.

It is easy to foresee some people objecting to such a specific definition and asking why we would want to limit the meaning of performance auditing when, for example, we have multiple interpretations of what evaluation involves. The distinction is that, unlike audit, evaluation as a form of social praxis cannot demand a specific legal position or evaluators benefit from specific procedures which guarantee that their reports are handled formally by parliament and other bodies, or that they have a legally regulated right to demand information. Auditing is defined within a certain context and concerns specific relationships. It also demands certain forms of freedom or independence. Without these prerequisites we cannot define an investigative activity as auditing. And if you, and this is certainly a normative statement, demand certain forms of independence and legal protection in order to do your work, it is a democratic necessity that you can define what you are doing.

The standpoint that the rationale of performance auditing as well as other forms of auditing has to be placed within a perspective of accountability has consequences for how we regard the role of auditing. An effect of reports issued by an SAI *can* be learning, but this has to be clearly distinguished from what can be seen as the rationale of performance auditing. It is also important to point out that this accountability perspective implies certain demands or restrictions on the information produced by performance auditors. The work has to be sufficiently reliable so that it can be used in the realm of accountability. This, in turn, tends to limit the questions that are dealt with by performance auditors; for example, about the long term effects of complex interventions.

The relationship to politics can also be seen as ambiguous. Within the SAIs the idea of a borderline between politics and policy objectives on the one hand and the implementation of policies on the other is regarded as fundamental, and is one that suits audit institutions. Yet the political debate is often about *how* different goals should be reached. For example, to discuss and suggest cuts or increases in spending programmes from a non-political perspective is, to say the least, awkward. So, it is easy to see that audit institutions face options, none of which can be regarded as satisfactory. To argue that political decisions are a no-go area for performance auditing means that the work will sometimes be very limited in scope and will be reduced to the scrutiny of technical implementations. To argue for the opposite raises issues of objectivity, and suggests that the auditor has a political role.

Finally, it is important to note that this chapter deals with performance auditing as it can be constructed as an ideal type. What we have described as the *raison d'être* of auditing as well as our definition of performance auditing does not mean that performance auditing cannot contribute to more operational improvements of programmes and implementation. In their actual work performance auditors can find things which are important for learning which will hopefully lead to learning and improvement, a subject covered by Lonsdale and Bechberger later in this book. Yet these are side-effects – important as they are – and in discussing the actual praxis of performance auditing institutions, should not be confused with the rationale for performance auditing.

NOTES

1. We will discuss and seek to define 'performance auditing'. In doing so we employ a term which is commonly used in international settings within organizations such as OECD, INTOSAI, EUROSAL, AFROSAL and the World Bank. However, it has to be said,

that at a linguistic level, performance auditing is a term which presents some difficulties. When translated and used in some countries it can have other connotations. The Swedish term *effektivitetsrevision* focuses, for example, on certain aspects of performance (efficiency and effectiveness). The French term *gestion*, used by the Cour de Comptes, seems to have other connotations more oriented towards management than the word 'performance'. And even several English-speaking Supreme Audit institutions use, or have used, terms other than performance auditing (e.g. value-for-money audit) with at least slightly different meanings. For a further discussion, see the introductory chapter of this book.

2. To avoid any misunderstanding, it is worth pointing out that the author of this chapter also can be said to have one leg in each camp.
3. The purpose of this chapter is not to give an overview of the development of performance auditing in different institutional settings. However, it is important to emphasize that the trajectories of the development of performance auditing in different countries differ in several respects, as do individual organizations' own views about the developments. It can be illustrated by a comparison between the United States and United Kingdom. GAO seems to have regarded evaluation not as part of auditing but as something complementary to it. It was not, in other words, defined as part of audit, but a way to develop programme evaluation on the federal level (The Oral History Project, 2009). In the case of the United Kingdom, value-for-money audit developed separately from evaluation, and more as a development or widening of traditional/financial audit (Pollitt et al., 1999, 233).
4. The following effort to grasp the meaning of evaluation is based on Furubo & Karlsson-Vestman, 2010.

REFERENCES

- Ahlbäck, Shirin (1999), *Att kontrollera staten. Den statliga revisionens roll i den parlamentariska demokrati*, Uppsala.
- Algemene Rekenkamer (2010), 'The Court of Audit's tasks', available at <http://www.courtfaudit.com/english/Organisation> (accessed 4 March 2010).
- Barzelay, Michael (1996), 'Performance auditing and the New Public Management: changing roles and strategies of central audit institutions' in *Performance Auditing and the Modernisation of Government*, Paris: OECD.
- Behn, Robert (2001), *Rethinking Democratic Accountability*, Washington, DC: Brookings Institution Press.
- Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale and Burt Perrin (2007), *Making Accountability Work*, New Brunswick, NJ: Transaction Publishers.
- Brante, Thomas, Heine Andersson and Olav Korsnes (eds) (2001), *Sociologiskt lexikon*, Stockholm: Universitetsförlaget.
- Bundesrechnungshof (2010), 'Audit rules of the Bundesrechnungshof', available at <http://bundesrechnungshof.de/we-about-us/legal-authority/audit-rules.pdf> (accessed 4 March 2010).
- Chelimsky, Eleanor (1995), 'Where we stand today in the practice of evaluation: some reflections', *Knowledge and Policy*, 8 (3), 8–19.
- Donaldson, Stewart I, Christina A. Christie and Melvin M Mark (2009), *What Counts As Credible Evidence in Applied Research?*, Thousand Oaks, CA: Sage Publications, Inc.
- EUROSAI (2005), *Public Audit in Europe. Achievements and Outlook*, Moscow: Accounts Chamber of the Russian Federation.
- Finer, Herman (1965), 'Administrative responsibility in democratic government'

- in Francis E Rourke (ed.), *Bureaucratic Power in National Politics*, Boston: Little Brown and Company.
- Flint, David (1988), *Philosophy and Principles of Auditing – An Introduction*, London: Macmillan.
- Friedrich, Carl J (1965), 'Public policy and the nature of administrative responsibility', in Francis E Rourke (ed.), *Bureaucratic Power in National Politics*, Boston: Little Brown and Company.
- Furubo, Jan-Eric and Ove Karlsson-Vestman (2010/forthcoming), 'Evaluation: for public good or professional power?' in Pearl Eliadis, Jan-Eric Furubo and Steve Jacobs (eds), *Evaluation: Seeking Truth or Power?*, New Brunswick, NJ: Transaction Publishers.
- Furubo, Jan-Eric, Rolf Sandahl and Ray Rist (eds) (2002), *International Atlas of Evaluation*, New Brunswick, NJ: Transaction Publishers.
- Gray, Andrew, Bill Jenkins and Bob Segsworth (eds) (1993), *Budgeting, Auditing and Evaluation*, New Brunswick, NJ: Transaction Publishers.
- INTOSAI (2004), *Implementation Guidelines for Performance Auditing*, International Organisation of Supreme Audit Institutions.
- Keen, Justin (1999), 'On the nature of audit judgements: the case of value for money studies', *Public Administration*, **77** (2), 509–25.
- Lane, Jan-Erik (2000), *New Public Management*, London: Routledge.
- Leeuw, Frans (1996), 'Performance auditing, new public management and performance improvement: questions and challenges' in *Performance Auditing and the Modernisation of Government*, Paris: OECD.
- Ling, Tom (2007), 'New wine in old bottles?' When audit, accountability and evaluation meet' in Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (eds), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Brunswick, NJ: Transaction Publishers.
- Lonsdale, J (2008), 'Balancing independence and responsiveness', *Evaluation*, **14** (2), 227–48.
- Lonsdale, Jeremy and Marie-Louise Bemelmans-Videc (2007), 'Accountability – the challenges for two professions' in Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (eds), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Brunswick, NJ: Transaction Publishers.
- Mark, Melvin M, Gary T Henry and George Julnes (2000), *Evaluation: An Integrated Framework for Understanding. Guiding and Improving Policies and Programs*, San Francisco: Jossey-Bass.
- Mayne, J (2006), 'Audit and evaluation in public management: challenges, reforms and different roles', *The Canadian Journal of Program Evaluation*, **21** (1), 11–45.
- Mayne, John, Marie-Louise Bemelmans-Videc, Joe Hudson and Ross Connor (1992), *Advancing Public Policy Evaluation*, Amsterdam: North-Holland.
- Mill, John Stuart (1886), *System of Logic Ratiocinative and Inductive. Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, 8th Edition, London: Longman, Green, and Co.
- Morin, D (2001), 'Influence of value for money audit on public administration: looking beyond appearances', *Financial Accountability and Management*, **17** (2), 99–118.
- Myrdal, Gunnar (1934), 'Finanspolitikens ekonomiska verkningar', SOU 1934:1 (Commission report published by the Swedish government).
- National Audit Office (2005), *State Audit in European Union*, London: National Audit Office.

- OECD (1996), *Performance Auditing and the Modernisation of Government*, Paris: OECD.
- Oral History Project Team (2009), 'The professional development of Eleanor Chelimsky', *American Journal of Evaluation*, **30** (2).
- Patton, M (2004), 'The roots of utilization-focused evaluation' in Marvin C Alkin (ed.), *Evaluation Roots, Tracing Theorists' Views and Influences* (pp. 276–92), London: Sage.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilka Summa and Marit Waerness (1999), *Performance Audit and Public Management in Five Countries*, Oxford and New York: Oxford University Press.
- Power, Michael (1997), *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press.
- Put, Vital (2005), 'Normen in performance audits van rekenkamers: een casestudie bij de Algemene Rekenkamer en het National Audit Office', Unpublished doctoral dissertation, Leuven, Belgium: University of Leuven.
- Riksrevisionen (2008), *Revisionens inriktning*, Memo within Riksrevisionen.
- Rist, Ray and Nicoletta Stame (eds) (2006), *From Studies to Streams*, New Brunswick, NJ: Transaction Publishers.
- Sartori, Giovanni (1970), 'Concept misformation in comparative politics', *The American Political Science Review*, **LXIV** (4).
- Schwartz, R (1999), 'Coping with the effectiveness dilemma: strategies adopted by state auditors', *International Review of Administrative Sciences*, **65** (4).
- Scriven, Michael (1991), *Evaluation Thesaurus*, 4th Edition, Newbury Park, CA: Sage Publications, Inc.
- Sharma, N (2007), 'Interactions and interrogations: negotiating and performing value for money reports', *Financial Accountability and Management*, **23** (3).
- Vedung, Evert (1997), *Public Policy and Program Evaluation*, London: Transaction Publishers.
- Weber, Max (1949), *The Methodology of the Social Sciences*, New York: The Free Press.
- Wisler, Carl (ed.) (1996), 'Evaluation and auditing: prospects for convergence' in *New Directions for Evaluation*, Number 71, San Francisco, CA: Jossey-Bass Publishers.

PART I

The conduct of performance audit

3. Selection of topics

Vital Put and Rudi Turksema

INTRODUCTION

Chapter 1 noted that the right to choose what to audit, how to audit, and subsequently what information to make public, is at the core of what gives Supreme Audit Institutions (SAIs) their unique position. It is a key part of the independence that Jan-Eric Furubo described as ‘fundamental’ to auditing. This chapter considers a key aspect of this unique position – the selection of performance audit topics. SAIs’ ability to select independently the topics for performance audits enables them to ask inconvenient questions to those in power, to focus on issues they consider of importance, and it also ensures that SAIs are ‘political’ players, in the sense that their work helps to influence decisions about resources (Lonsdale, 2010). At the same time, study selection is one of the most challenging aspects of auditing. It is often not clear what audit clients expect from auditors or whether stakeholders, and those reading the reports, find performance audit reports useful or are able to use the information and recommendations contained within them.

In the field of evaluation, Weiss (1998) and Patton (1997) have suggested that, from the viewpoint of effectiveness, it helps if evaluators give significant weight to the considerations of their stakeholders when undertaking their work. On the other hand, too much responsiveness raises questions about independence and might negatively affect the credibility of those doing the work. This is particularly the case for auditors and as a result, the issue of what to audit turns into a balancing act between listening to stakeholders and safeguarding their independent position (see also Lonsdale, 2008; Van der Meer & Edelenbos, 2006; Wilkins & Lonsdale, 2007).¹

Closely connected to this issue, and important for the study selection process, is the expected effectiveness of proposed new performance audits. How their effectiveness is predicted and ultimately measured is dependent on how we frame or judge effectiveness. In the context of performance auditing ‘accountability’ and ‘learning’ are two important frames as

Chapter 1 makes clear (see also Bemelmans-Videc et al., 2007; Lonsdale, 1999, 2008; Van der Meer, 1999; Van der Meer & Edelenbos, 2006). On the one hand, we have the ‘responsiveness/learning frame’: an audit office that focuses on the actual performance of the public sector and is responsive in how it does its work, and one for whom government bodies might be considered as the main clients. On the other hand, we have the ‘accountability frame’, where parliament is the main client, the focus is much more on performance compliance and transparency, and where independent opinions are considered of prime importance. Here too, we see that it is difficult for audit offices to find a good balance between the two ambitions.

There are various good reasons why it is important that audit offices select the right topics for their performance audits. Firstly, like the public entities they audit, they use public money to perform their tasks. There is thus a very strong case that the same criteria they use to assess the work of others apply to audit offices themselves. Efficiency and effectiveness are probably the most important criteria. In terms of the performance-audit selection process this means that since the resources of SAIs are limited and there is at the same time a large ‘audit universe’ to be audited – SAIs should be looking to select those topics that they expect to be most cost-effective. This can be very difficult, as an SAI may have different perceptions of what constitutes effectiveness and it might struggle with the tension between independence and responsiveness. Whatever the conceptualization of effectiveness and the information analysed, there will be no scientific means of establishing an optimal programme of performance audits, and decision-making by the institution always involves a high level of judgement. Moreover, they cannot foresee the future: this year’s brilliant choice for examination may turn into next year’s worst decision as events in the environment they are auditing in move on.

Secondly, it is also important to take the opportunity costs of performance auditing into account when selecting topics for examination. A performance audit usually requires considerable time and attention on the part of the audited body, so it is important for auditors to recognize this and address it properly. If not, the auditor is at risk of delivering an audit that does not add value to its intended users and at the same time possibly distracts the administration and parliament from more important tasks. For example, when a government is focusing its policies primarily on finding a way out of the financial crisis, it will be probably not very receptive to auditors telling it to improve its *ex ante* analysis of possible government interventions to remedy such a crisis. Rather, in such circumstances, audit offices should perhaps focus their performance audit work on ways

for government to get the same value for less money. This line of argument suggests that performance audit topic selection should, by and large, be in line with the current political debate if auditors want to have influence.

The challenges associated with performance-audit topic selection outlined above help to explain why audit offices give the matter so much attention. In this chapter, we examine how a number of SAIs – primarily the Rekenhof in Belgium and the Algemene Rekenkamer in the Netherlands – deal with these issues and how their decisions affect the selection of performance-audit topics.

A THEORETICAL APPROACH TO THE SELECTION PROCESS

Previous Research

There are few academic studies of how SAIs pick out topics. In their comparison of performance auditing in five countries (Finland, France, the Netherlands, Sweden, United Kingdom) Pollitt et al. did consider the subject (1999, 109–112), noting the following.

- The selection of topics was an annual exercise.
- Five selection factors were influential in all five countries: the volume of public resources involved in an activity, the likely risk to public funds, systematic coverage of the audit field, the follow up of topics where earlier audits demonstrated there were problems, and topics that are of strong political salience. The relative weighting of these criteria can vary from time to time and from SAI to SAI. These criteria are not determinative and still leave plenty of room for debate and judgement.
- Topic selection was predominantly a bottom-up process (ideas rise up from performance auditors), but more recently strategic guidance from the head of an SAI or from senior staff had become more important.

In a Sigma study on the relations between SAIs and parliamentary committees in Central and Eastern European countries (Sigma, 2002, 22–3), the influence of parliamentary committees on the SAIs' audit programme is discussed. Its findings include that:

- The major part of work carried out by SAIs depends upon the SAIs' priorities and decisions.

- The influence of parliament and parliamentary committees on an SAI's audit programme may take the form of their right to review or approve the annual audit programme, or to request additional audits.
- Apart from this there is an informal practice of parliamentary committees (or even individual members of parliament) submitting suggestions on audit subjects. They may be included in the SAI's audit programme at the discretion of the head of the SAI.
- There is no evidence in any of the countries suggesting that parliaments or parliamentary committees prevent audits from being planned or performed by the SAI.

As part of a comparison of performance auditing practices in three countries (Canada, the Netherlands and the United Kingdom), Sterck et al. also looked at the selection process (2006, 33–5, 62–4, 92–5). They found that in all three countries the selection process was based on:

- a monitoring system, that is, a systematic collection of information on the audit universe (parliamentary documents, strategic plans, information coming from personal contacts, et cetera); and
- a mix of criteria (for example risks, finances involved, feasibility) and strategic guidance (for example prioritizing certain policy areas).

O'Connell (2007) made a study of the relationship between the US Congress and the Government Accountability Office (GAO). She examined whether the GAO chose to audit policy programmes as a 'neutral watchdog for waste' or as a 'political' auditor² that sought to advance particular policy objectives. In O'Connell's model, an auditor can be a Democrat, a Republican or non-partisan, and government programmes are either Democrat (education, health, labour, housing) or Republican (defence, home security, commerce). Non-partisan auditors wish to maximize the amount of waste they report from an investigation, partisan auditors wish to maximize the amount of reported waste from projects affiliated with the opposing party and to minimize the amount of reported waste from projects linked to their own party. Building on game theory she elaborates a series of assumptions which are subsequently (partially) empirically tested with data on investigations by the GAO from 1978 to 1998. She concluded that the GAO was indeed a neutral watchdog.

Lonsdale (2011) distinguishes four models of study selection. These are derived from the two dimensions already discussed – independence and responsiveness – that are, according to Lonsdale, characteristic of how SAIs choose their performance-audit topics. Taking this further,

he distinguishes between ‘unrestrained’ and ‘constrained’ independence. Some SAIs, such as the UK NAO, rely largely on the auditor’s knowledge and professional judgement and do not exclude particular areas of government before developing a programme of work (Model 1. *Auditor judgement and prioritization*). Other SAIs, like the Algemene Rekenkamer, have developed more deliberate strategies, with a significant focus on some key areas (Model 2. *Deliberate strategies*). In the second dimension, Lonsdale distinguishes between ‘reactive’ and ‘proactive’ responsiveness. In the former, an SAI is very open to relevant outside requests (Model 3. *Responses to external requests*). The GAO is an example of such an SAI. In the final model (Model 4. *Taking account of consultation exercises*) the SAI takes a more active role and consults with its stakeholders when preparing their activity programme.

DIMENSIONS TO CHARACTERIZE THE SELECTION PROCESS

From the studies we have described and also from our own experience, we can distil some dimensions that can characterize an SAI’s selection process. We distinguish three broad categories: institutional, organizational, and operational.

Institutional Level: Position in State Structure, Audit Mandate

The institutional position of an SAI in a country’s state structure, and how this position – in terms of finding a balance between responsiveness and independence – is interpreted by the SAI, is an important factor for its selection process. The GAO, for example, states explicitly that it ‘works for Congress’ (GAO, 2009a), whereas other SAIs are associated with their parliament, or are positioned independent of both parliament and government (for example in the Netherlands and the United Kingdom). These institutional arrangements have immediate consequences for the selection process of an SAI. The audits of the GAO are predominantly (about 90 per cent) at the request of Congress³ (Azuma, 2004, 91; GAO, 2003, 1–2), whereas requests from parliament make up just a small proportion of the audit portfolio of the Rekenhof and the Algemene Rekenkamer.

In addition, the audit mandate or remit is an important factor to take into account. The audit mandate varies very much between SAIs. Some audit only at the level of central government and its associated public bodies, whereas others are responsible for auditing all government levels, as well as state enterprises and a range of other public bodies (NAO, 2005,

14–18). It seems likely that an SAI that has to audit multiple layers of government will make different choices in its selection process compared with an SAI that only has to audit central government.

Organizational Level: Mission, Strategy, Resources

To guide their work SAIs often set out an explicit mission for themselves. For many, this has two parts: to improve both the performance and the accountability of the public sector as Jan-Eric Furubo highlighted in the previous chapter. Many audit offices have stated in their mission that – through their performance audits – they want to improve public sector performance. Typically, their mission statements use phrases such as: ‘focus on effectiveness’, ‘create value’, ‘share knowledge of good practices’ (Danish Audit Office, 2008). At the same time, audit offices want to provide their parliaments with objective and audited evidence on the performance of government bodies.

Secondly, many SAIs have developed a strategic plan in which they describe their goals and how they will achieve them. In these plans, SAIs set out explicit choices with regard to what to audit. For example, the Dutch Algemene Rekenkamer’s strategic plan 2004–09 focuses its performance audits on areas where a weak link between policy and policy implementation has consequences for public services, security, and the environment. Other SAIs such as the Belgian Rekenhof, put more emphasis on the systematic and periodic coverage of the audit field.

A third, and very restrictive factor at the organizational level, is the availability of resources. The ‘audit universe’ (the full set of organizations that an SAI can audit) is very extensive and there are many potential topics for audit, but resources are limited. The Algemene Rekenkamer, for example, has to audit €350 billion of public expenditure with about 80 performance auditors (Turksema & Van der Knaap, 2006). It is then a significant challenge to choose the topics for performance audits that contribute most to an SAI’s mission.

Operational Level: Monitoring, Risk Analysis, Selection Criteria

Within the restrictions created by the factors at both the institutional and organizational level, SAIs have to make choices on how to select topics for performance audit. In many SAIs the selection process is nourished by a monitoring system: they systematically collect information on the departments, public bodies, or other organizations they can audit. This encompasses written documentation (policy plans, strategic plans, parliamentary documents, statistics, evaluations, news coverage, reports of the

ombudsman, etc.), as well as softer data, such as inside information from personal contact with civil servants.

In most SAIs, some sort of risk analysis is applied to guide the selection process, drawing on their accumulated knowledge. Often this takes the form of an analysis – as objective as possible – of the chance that an event will occur, linked to the potential impact of this event. The end result of such an SAI-wide risk analysis is – or should at least be – an objective ranking of proposals. As it is in reality impossible to make fully objective assessments of risks, the assessment is largely based on field knowledge and the auditor's professional judgement. As a result, there will be certain subjective elements to this judgement.

In such risk analysis, selection criteria have to be used. They can be derived from, for example, the SAI's mandate, mission, and/or strategy. Popular selection criteria are the presence of risks and/or problems related to effectiveness and/or efficiency, the financial significance of a topic or its importance for citizens, the likely added value from doing the work, signs of interest on the part of parliament, the feasibility of an audit, and the desirability of balanced coverage across the audit universe. Both the NAO in the United Kingdom and the GAO in the United States pay special attention to the potential financial impact of an audit, as they have set themselves specific goals on the rates of return from their work, including performance audits. The NAO aims to save £9 for every £1 it costs and in 2008 managed 11 times its cost (NAO, 2009). The GAO claims – albeit on a different basis – to have returned \$114 for every dollar they have cost in 2008 (GAO, 2009a).

The GAO takes a somewhat different approach with its risk analysis, which it issues as a separate report. This takes the form of a *High Risk List* (GAO, 2009b) in which it describes which government operations it has designated as 'high risk'. To determine which operations should be designated in this way, the GAO has established a methodology (GAO, 2000). The GAO provides its high-risk reports at the start of each new Congress, and in its 2009 update for the 111th Congress, GAO identified 30 high-risk areas (GAO, 2009b). If significant progress is made in correcting the risks, areas are removed from the high-risk list.

Finally, certain external factors at the operational level may affect the selection of performance audit topics, for example, specific requests by the parliament. Being in most cases their primary client, SAIs are usually very responsive to requests from their parliaments, committees within them or, in some cases, individual members of the legislature. The UK NAO, for example, presents its draft programme to the Committee of Public Accounts annually in seeking comments and inviting suggestions. In addition, unsatisfactory progress in a department or policy area where

an SAI has previously done audit work may be a good reason for the audit office to do a follow-up audit. Moreover, collaborations with other SAIs, universities, or Inspectorates may also lead to the identification of other audit proposals.

OTHER FORCES THAT SHAPE THE SELECTION OF AUDIT TOPICS

In the previous section, we described the more ‘formal’ factors that determine the selection of performance-audit topics. In addition to these, there are also more ‘informal’ factors that play a role.⁴ Firstly, the professional (personal) judgement of auditors and others involved in the selection process plays a very important role.⁵ It is not possible to know all of the potential risks, and the limited resources of SAIs mean that some audit areas may seldom be covered. Thus, the knowledge, experience and judgement of individuals play a role. In itself, there is nothing wrong with what we might call professional judgement, as long as personal preferences do not result in ‘thinking backwards’ (selecting first, and then gathering evidence to justify this choice).

Secondly, political taboos can play a role. This is also acknowledged in the INTOSAI implementation guidelines which state that: ‘Some SAIs may choose topics based on strategic choices rather than selection criteria (for example, with regard to the type of performance audit, policy spheres, relationship with reforms within the public sector, and so on). Sometimes these strategic choices may reflect the constitutional and legal conditions and the established traditions. They may also reflect “political realities” (that is, certain topics are not expected to be subjected to auditing)’ (INTOSAI, 2004, 45). Again, there is nothing wrong with this, as long as it is clear that these choices are made for justifiable reasons. Audit institutions may also be risk averse (in seeking to avoid controversy and confrontation) or risk seeking (to perhaps become better known), with this influencing the selection of topics.

A third informal factor is the organizational structure of an SAI. Many SAIs are organized along the lines of certain policy areas (for example departments) and/or along topical areas (performance audit, financial audit). Auditors working in such units have specific expertise in some audit areas and build up knowledge in those fields. This seriously limits the possibility of working with a fully risk analysis-based selection procedure, because it would imply that auditors are fully interchangeable. That being said, some SAIs – the UK NAO, for example – have introduced more flexible, matrix-style working arrangements and do expect their auditors to

be able to move between sectors and apply their technical knowledge in different settings.

Finally, SAI publishing policy and reporting practices will also influence the selection process. For example, if an SAI only publishes extensive, broadly focused reports, only areas where there is enough material to justify a fully fledged report will be selected. Similarly, the number of publications over a period will also determine if an area is covered and how frequently.

ILLUSTRATIVE CASES: THE SELECTION PROCESS IN BELGIUM AND THE NETHERLANDS

In the previous section we described the formal and informal factors that drive the selection of performance audits. In this section we present two cases in more detail: Belgium's *Rekenhof* and the *Algemene Rekenkamer* in the Netherlands. We base our analysis on documents from those two SAIs, such as strategic plans, annual reports and manuals.

BELGIUM

The Position of the Rekenhof

Belgium is a federal state, with a federal level, three regional levels (the Brussels, Flemish and Walloon regions) and three communities (French speaking, Dutch speaking, German speaking). The Belgian Rekenhof undertakes the audit of the Federal State, the Communities, the Regions, the public service institutions dependent upon them, and the provinces. The control of municipalities falls outside the scope of its powers.

The federal structure of the Belgian State is also reflected in the internal structure of the Rekenhof. There are both federal and regional directorates. Since these directorates work for other parliaments, they can highlight other issues and so they have separate strategic plans which guide the selection process.⁶ The Rekenhof has close links to the Chamber of Representatives. The Members of the Court are elected by them and the Rekenhof's budget is approved by the Chamber of Representatives.

Study Selection Process

The Belgian Court of Audit uses four criteria for selecting studies, which are incorporated into its mission statement: 'The selection of audit themes

is mainly carried out on the basis of a regular *risk analysis* and the *interests of the parliaments*. The audit themes are *spread in a balanced way among all aspects of government policy*. . . . *the work of others is used* . . .’ (Rekenhof, 2004, 12). This mission statement reveals itself in the following criteria:

1. ‘*risk analysis*’. The Rekenhof wants to contribute to a better functioning government. Therefore, performance audits are focused on issues where there is a *need* and a *scope* for improvement.
2. ‘*interests of the parliaments*’. As auditor of the parliament(s), the Rekenhof wants to select issues that will attract the attention of parliamentarians.
3. ‘*spread in a balanced way among all aspects of government policy*’. The Rekenhof strives for a balanced distribution of its performance audits over all policy areas. In practice, this means that there has to be at least one performance audit done in every policy area (defined as a department with the accompanying agencies), within a cycle of five years. This means that the risk analysis is carried out on a cyclical basis. (As we will see later, some areas – the tax system, the social security system and personnel expenditures – are excluded from this process because of their financial importance and are subject to several performance audits every year).
4. ‘*the work of others is used*’. The Rekenhof is not the only body that publishes reports about the performance of public organizations. In order to avoid duplicating the work of others, audits are focused on areas with identified information gaps. Adding value can involve generating new information, but can also mean producing better (for example, more reliable or profound) information. Added value is maximized when an audit tackles forgotten or unknown problems (so-called *Cinderella subjects*).

The Belgian Court of Audit’s risk analysis is based on a set of indicators. In an interrogative form, they can be formulated as follows:

1. Are there indications that the goals of a policy are not being achieved or that the outputs are unsatisfactory (for example poor quality of service, delays)?
2. Is the policy based on a clear policy theory (for example clear objectives, clear link between instruments and objectives)?
3. Is there a system of result-oriented management (for example strategic plans, performance measurement)?
4. Is there a clear organizational structure?
5. Is there a system to assure the quality of the operating processes?

6. Are opportunities to apply information technology sufficiently exploited?
7. Are basic human resource management (HRM) instruments applied (for example the work force is based on a recent evaluation of staffing needs, clear job descriptions)? Are there HRM-problems (for example key positions not occupied, high staff turnover, high absenteeism)?
8. Is there collaboration with other actors (public or private) where necessary?
9. Is there sufficient 'external pressure' to perform well (for example do clients have the chance to complain, is there competition, is there an obligation to give account, do supervisory organs play their role)?
10. Are there external changes that affect the performance of the public service (e.g. strong increase of the target group)?

The reasoning that underpins these indicators is: if a policy or a public service fails (indicator 1), this can be attributed to a poorly designed policy (indicator 2), to poor implementation (indicators 3–9) or to difficult external circumstances (indicator 10). In other words: 'bad results' = f ('bad policy' and/or 'bad implementation' and/or 'bad luck') (Figure 3.1).

The Belgian Court of Audit's audit manual (Rekenhof, 2006a) states that the first indicator is the most important one. It directly relates to outputs and the outcomes to be achieved, whilst the other indicators relate to conditions that are important for achieving these outputs and

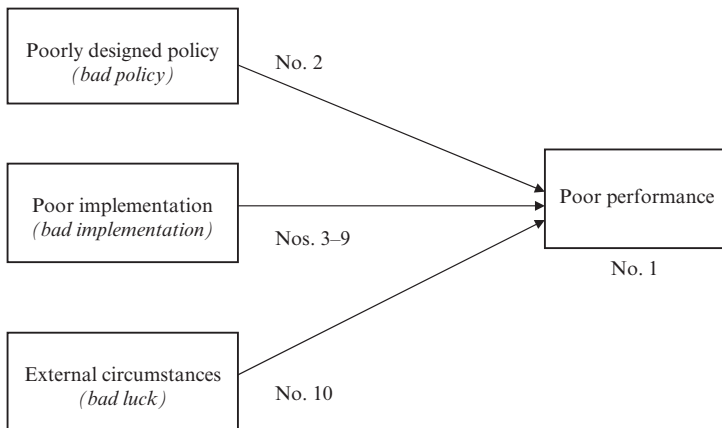


Figure 3.1 Logic that underpins the risk indicators of the Belgian Court of Audit

outcomes. The significance of the other indicators depends on the nature of the activity. For example, in public services that carry out repetitive activities (such as a tax administration, a social security administration), computerization (indicator 6) will be an important success factor and get a higher weight; in a probation service, the professionalism of staff (indicator 7) will probably be the most important factor. This assessment strongly depends on the judgement of the auditor.

Apart from this risk analysis, there are also some explicit and implicit strategic choices that influence the selection process within the Belgian Court of Audit. For example:

- Strategic plans such as the federal strategic plan (2006–09) (Rekenhof, 2006b) require more audits to be focused on outputs and/or goal achievement, and also recommend that more follow-up audits are carried out to determine if problems addressed in prior audits have been resolved.
- There are also more implicit strategic lines of force. For example, in the organizational structure of the Court of Audit there are separate departments for audits of the tax system and of the social security system. By creating a separate structure and by providing more staff at the disposal of these areas, they are implicitly prioritized.

The selection process in the Belgian Court of Audit is nourished by continuous monitoring of all policy areas and systematic collection of information on the departments, agencies and other bodies that it audits. This monitoring encompasses policy plans, strategic plans, parliamentary documents, statistics, evaluations, news coverage, reports of the ombudsman, and so on. Monitoring documents are stored in a highly structured electronic database (the ‘permanent file’). Based on these documents, annual monitoring reports are produced to underpin the selection process.

THE NETHERLANDS

Position of the Algemene Rekenkamer

The Kingdom of the Netherlands is a constitutional monarchy and a decentralized unitary state. As well as central government, there are 12 provinces and about 400 municipalities which have some executive power and their own elections. The Algemene Rekenkamer’s audit remit is limited to the central government. Provinces and municipalities have

their own audit functions (as highlighted by Katrien Weets' later chapter in this volume on the Rotterdam Rekenkamer). The Members of the Rekenkamer's Board are appointed by the government, on the basis of recommendations by Parliament.

Strategic Choices

With the Strategy 2004–09 (Algemene Rekenkamer, 2004b), and the subsequent Strategy 2010–15 (Algemene Rekenkamer, 2010), the Algemene Rekenkamer applied a new approach to the selection of audit topics. Three characteristics describe this approach best: (1) more emphasis on recurring audits in order to identify trends; (2) more emphasis on effective influencing; (3) focus on a limited number of policy areas.

Recurring audits

With its Strategy 2004–09 the Algemene Rekenkamer puts more emphasis on recurring (or follow-up) audits, arguing that a single audit is not as effective as a series. This is in line with its ambition to shift the focus from identifying 'events', based on individual audits, towards identifying 'trends' that can provide insight into underlying structural opportunities to improve the performance of government. In the Algemene Rekenkamer's view, SAIs are in a unique position to point out important trends and structures, much more than they are able to effectively respond to short-term issues.

Effective influencing

In addition, the Algemene Rekenkamer pays increasing attention to the effective influencing of its stakeholders. This implies that the audit report is not considered the end product of its audit work, but rather an intermediate step to be used in further and continuing discussion with its stakeholders. This can take the form of, amongst other things, expert meetings, conferences, journal articles, or presentations of the outcomes of risk analyses to the auditee.

Limited number of policy areas

The Strategy 2010–15 gives priority to those policy areas (domains) where the Algemene Rekenkamer believes a weak link between policy and implementation has the greatest consequences. These are:

1. 'People': healthcare, education, development aid.
2. 'The Netherlands': competitiveness, integration, infrastructure, security.
3. 'Planet': climate, energy.

Within these policy areas the Algemene Rekenkamer starts from the analysis of societal problems, with the aim of gaining insight into policy and implementation (Algemene Rekenkamer, 2004a, 3). A further selection is made with the help of seven selection criteria:

- Complex policy chains (e.g. when several ministries or several local actors are involved).
- Financial value.
- International treaty obligations.
- Use of unique powers.
- Insight thanks to helicopter view.
- Societal urgency.
- Opportunities for improvement.⁷

Selection Process

These strategic choices obviously limit the possibilities for coming up with new audit topics in the annual selection process. However, the Algemene Rekenkamer devotes some attention to this, as it is closely connected to the monitoring process, which also serves as an instrument in its influencing strategy and management of client relationships. Adequate monitoring remains important because it enables the organization to validate the strategy, offers the opportunity to discuss topics with government bodies that are not part of the strategy, and may lead to some interesting audit proposals.

For its annual working plan the Algemene Rekenkamer uses three inter-related tools: monitoring, issue matrices and integrated risk analysis. The relationship between these tools is shown in Figure 3.2 below.

Monitoring is a standard activity carried out by all of the Algemene Rekenkamer's organizational units. In each case, the ministries' policies and the performance of RWTs (rechtspersonen met een wettelijk taak – or legal entities with statutory tasks) are monitored. Given the strategic ambition of helping to reduce social problems, the Algemene Rekenkamer is also increasingly monitoring social developments.

The *issue matrix*, the second tool, is an outcome of the monitoring and is intentionally designed to facilitate discussion of the monitoring findings with the Algemene Rekenkamer's Board. Issues are summarized in memos that answer the following four questions: (1) What is at issue? (2) Is it undesirable? (3) Where does it occur? and (4) Who are the main players? The issue matrix is used to identify relevant issues that are occurring in the Algemene Rekenkamer's audit field but are not necessarily covered by its strategy.

The third tool is *integrated risk analysis*. In contrast to monitoring and issue matrices, risk analysis builds on the strategy and is more focused so

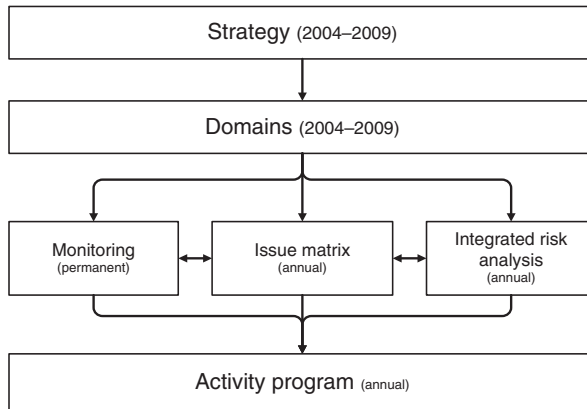


Figure 3.2 *Relationship between the strategy and the planning and selection tools*

that it can be used directly for the Annual Work Plan and the planning of the regularity audit. Integrated risk analysis is a systematic and efficient means to generate, analyse and record information on the entire audit field and so identify and classify the risks in the domains that are relevant to the Algemene Rekenkamer. The risk analyses are integrated in that they:

- combine all identified and prioritized risks so that choices can be made for both the regularity audit and the annual programming of performance audits; and
- are conducive to the detection of connections between the operation of public administration and risks to the performance of public administration.

The system used to assess risk is very similar to that used for the issue matrices. The risk is estimated as a combination of the likelihood of an undesirable event or situation occurring and its impact on one of the risk domains: $\text{risk} = \{\text{likelihood of risk} \times \text{impact of risk}\}$. The Algemene Rekenkamer uses the model shown in Figure 3.3.

The end product of integrated risk analysis is a systematic overview of the main risks to the operation and performance of public administration and associated third parties. The analysis results serve as input into proposals for regularity audit (selection of auditee and approach) and the annual programming of performance audits. Integrated risk analysis also contributes to the accumulation of field-specific knowledge and to the exchange of this knowledge among organizational units.

Probable impact	Significant	Average	High	High
	Average	Low	Average	High
	Insignificant	Low	Low	Average
		Improbable	Possible	Probable
	Likelihood of occurrence			

Figure 3.3 Risk matrix

The end result of the Algemene Rekenkamer's annual selection process is a list of audit topics that consists of a number of recurring audits (for example, the European Union Trend report and the State of the Policy Information), a number of audits resulting from the process described above, and a number of audits requested by parliament. The list of audit topics can be found in the Annual Work Plan (Algemene Rekenkamer, 2008), which is published on their website and sent to parliament.

A CRITICAL APPROACH

SAIs' attempts to try to make the selection process based on objective criteria is a positive step, but despite this the selection remains a subjective process, requiring judgement. In this section we will argue that selection of topics is an inherently political (in the sense of value laden) activity, and that, for the time being, there is little evidence to support the premise that risk analysis is a valid method for selection of audit topics.

SELECTION IS A VALUE-LADEN PROCESS

In considering evaluation, Weiss states that there is an implicit political statement in the selection of some programmes to undergo evaluation, while others go unexamined (Weiss, 1987, 47; for similar considerations, see: Bovens et al., 2006; Chelimsky, 2007; House & Howe, 1999; House, 2006, 121). The same can be said for performance audit. For example:

- Strategic choices, such as giving priority to certain policy areas or societal problems, are evidently value-laden choices. SAIs can have

their ‘own agenda’, and these priorities can coincide with or can counterbalance the policy of the government of the day.

- The selection of topics is a way to steer the political agenda and debate. It is not only the selection of issues and problems which should get attention in the public debate, but also the language in which a problem is framed, which suggests some causes, solutions and recommendations, and therefore pushes the political debate in a certain direction (see FrameWorks Institute, 2002). For example, a ‘capacity problem in the prison system’ can be framed as an ‘over-population problem’ or it can also be labelled as a ‘lack of cells’. Both labels suggest different causes and solutions. The drug problem can be defined as a ‘crime problem’, but also a ‘medical problem’ or an ‘economical problem’. All these frames look for different causes and solutions. Each point of view will lead to another type of assessment.
- The share of audits in the audit portfolio of an SAI that have been requested also has a political dimension. As we noted at the start of the chapter, an essential element of the independence of an SAI is the freedom to choose topics for audits. This is also incorporated in the Lima Declaration, section 13.1: ‘Supreme Audit Institutions shall audit in accordance with a self-determined programme. The rights of certain public bodies to request a specific audit shall remain unaffected’ (see also: Mazur, 1994, 13–15). In addition, the INTOSAI manual on performance auditing states: ‘The SAI must maintain its political neutrality, but maintenance of the SAI’s independence does not preclude requests to the SAI from the executive, proposing matters for audit. However, if it is to enjoy adequate independence, the SAI must be able to decline any such request’ (INTOSAI, 2004, 44). A high proportion of requested audits could crowd out the own-initiated programme of the SAI, and in that way diminish its independence. Much depends on the arrangements, for example: Is there separate funding for audits on request? Who determines the terms of reference? Can the SAI refuse a request?
- Very often SAIs take into account the ‘financial importance’ of a subject. This is also not a neutral choice. Apart from the fact that this assessment is influenced by the available financial information (the structure of the budget, accounting system, etc), it means that government policies that might have a big impact on citizens but not involve substantial costs (such as rent acts, matrimonial law), are likely to be ignored.
- Seemingly neutral tools, such as risk analysis, are also not value-free. The design of a risk analysis system implies several value-laden choices (Raaum & Morgan, 2001, 126). For example:

- the choice of the entity for analysis, the audit unit, is not neutral. Focusing the detailed risk analysis at process level (e.g. functional processes such as procurement) as opposed to broad risk analyses that focus on the entire government system and policy issues, will lead to different types of audits; and
- the choice of risk indicators can include managerial risk indicators that concern internal processes as opposed to risk indicators that concern policy issues (e.g. quality of policy objectives, quality of regulations).

If we combine both parameters, we can distinguish two extreme design options. The first might be a risk-analysis system focused on processes, using only managerial indicators. This is likely to result in management audits, in which only civil servants are blamed. In contrast, a risk analysis that starts from problems in society, using indicators focused on policy issues, is more likely to result in more socially relevant, and consequently more politically sensitive, themes, where ministers have to account for their policy.

THE QUESTIONABLE EVIDENCE BASE OF RISK ANALYSIS

‘Risk analysis’ is the overriding mantra in the discourse about study selection and it therefore deserves further critical examination. We will first discuss some basic concepts, starting from the steps required to design and implement a risk-analysis system. A note of caution is needed to avoid misunderstandings. Firstly, the comments below only concern the application of risk analysis in the context of topic selection for performance audits within SAIs. They should not be generalized to other contexts. Secondly, a ‘risk’ is a probability that something goes wrong, a ‘problem’ is something that actually goes wrong, such as insufficient quality of service, waste. Very often systems of risk analysis of SAIs do not only consider ‘risks’ but also ‘problems’. The comments below mainly relate to risks (but some of them are also relevant for problem analysis).

Raaum and Morgan (2001, 126–36) identify seven steps (the first four concern the design of a risk-analysis system, the next steps relate to the implementation of this system). They are:

1. *Define the auditable units.* The ‘audit universe’ has to be broken down into units against which risk assessment will be applied (for example organizations, programmes, expense category). This results in a list of auditable units.

2. *Choose selection factors.* Selection factors are criteria for identifying issues such as financial significance, knowledge of problems, risks, public interest, recent audit coverage. The auditable units are the rows in a matrix, and the selection factors are columns in the matrix.
3. *Select weights for the factors.* Audit offices can treat each selection factor equally, or they can also assign them different weights.
4. *Establish rating scales for the factors.* The auditor could opt for a numerical scale (for example 1 to 5) or a descriptive scale (for example high, medium, and low).
5. *Obtain data on chosen selection factors.* Common information sources are strategic plans, performance reports, newspaper reports, input from management officials, and so on.
6. *Rate the auditable units and calculate the final score.* Using the data collected on each factor, scores are assigned for each factor to each auditable unit. For each auditable unit a total score is calculated.
7. *Select and prioritize audit selections.* The assessment results in a list of potential auditable audits,⁸ ranked from the highest score to the lowest. In making the final selection, other factors will be considered, such as the availability of audit resources.

Risk analysis is a technique that is supposed to detect (potentially) failing public organizations, policies and functions. The validity of risk analysis has been neither confirmed nor invalidated empirically. There is no proof that risk analysis is indeed capable of maximizing true positive choices (for example choosing failing organizations) and minimizing false positives and false negatives. The usefulness of risk analysis as a selection technique for studies depends on the accuracy of assumptions upon which it is based and which are challengeable and not exhaustive.⁹ The assumptions include:

- The assumption that risks in different policy areas can be compared. This is highly questionable. It is difficult to argue that risks in, say, the public health area are more or less important than those related to environmental issues, education or defence. There is no indisputable common denominator with which to compare these areas.
- The assumption that risks are unequally spread over audit units. Without this, risk analysis is pointless but a closer look at some of the risk indicators used such as ‘complexity’, and ‘number of actors involved’ makes it clear that a lot of risk indicators have no selective value, with almost the entire public sector at risk.
- The assumption that risks are stable as time goes by. This is not unimportant since there can be a long lapse of time between the

selection of an audit topic and the publication of the resultant audit report.

- The assumption that risks are knowable, in other words, that it is possible to collect complete and reliable information about the probability and impact of all sorts of internal and external risks that can hamper good management and goal achievement. Yet auditors very often lack subject knowledge, which could result in relying on available risk analyses of departments and agencies. This could increase the probability that auditees who share an honest and well elaborated risk analysis would be selected for an audit ('false positive' decision error). In addition, even if auditors do possess the necessary subject knowledge, there will be incalculable risks which are inherently unknowable.
- The assumption that there are clear policy and managerial objectives. Risk is the likelihood that the organization will fail to fulfil its objectives. Risk analysis involves systematically identifying and assessing risks that might affect the achievement of objectives. If objectives are unclear, implicit, contradictory, outdated, and so on, then identifying risks will become difficult.
- The assumption that it is organizationally possible to reshuffle staff every year, depending on the results of the risk analysis. For example, if ICT appears to be an important area of risk in a certain period, then the SAI should have the flexibility to acquire the necessary knowledge and skills. This cannot be guaranteed or may take time.

Leaving aside these problems, implementing risk analysis requires substantial investment, in people as well as in information technology. There is no information on the cost-effectiveness of this selection method in comparison with other selection methods such as problem analysis or indeed, random selection.

CONCLUSION

Although SAIs approach the selection process in various ways, there are obvious similarities. They all base their selection process on a system of monitoring, all use some form of risk analysis, and all allow strategic choices to guide the selection process. Requests from the legislatures also play a role. The exact shape of these instruments and the relative weight of each one can be very different; for example, for the Rekenhof, requests from the parliament(s) are rare. The selection process is mainly based on

the application of a number of criteria. To a lesser extent, strategic options also play a role. Within the Dutch Algemene Rekenkamer strategic choices are of greater weight, and risk analysis is (primarily) used within this strategic framework. Furthermore, each year several performance audits are carried out on request of parliament.

Generally, SAIs provide information to the public and stakeholders on their approach to topic selection, so are transparent and can be judged on this basis as well as on the programme of performance audit reports. However, none clearly articulates how the effectiveness of the selection process could be evaluated. At best one can look at the effectiveness of the overall performance audit programme to look at the combined results of topic selection, conduct and reporting of performance audits.

There is a need for more research to describe how decisions on topic selection are made and to gain insight into the individual, organizational and external factors that influence this decision-making process. Purely document-based research like this is not sufficient as SAIs are likely to rationalize and simplify their selection processes when they describe them in formal publications. Research techniques like participative observation or in-depth interviews with all ranks in an SAI are more valid approaches to get behind this 'formal front'.

Given our current state of knowledge, the question of whether some selection systems are more effective than others cannot be answered. Research is needed to test the validity of different approaches, such as how well does risk analysis perform in comparison with, say, problem analysis or the random selection of topics. We have also demonstrated that the seeming neutrality and scientific underpinning of risk analysis is, at least partly, a *trompe-l'oeil*. Beneath the seemingly exact, almost actuarial, assessments, lie values and unproven assumptions.

NOTES

1. This tension between independence and responsiveness is especially characteristic of performance auditing, as there are no strict rules and regulations to guide the audit work, as compared to financial auditing.
2. O'Connell uses the word 'political' in a specific way to mean 'partisan', whereas we will use it later on in this paper in the broader sense of value-laden.
3. This was not always the case. For example, in 1969 only 10 per cent of the GAO reports were carried out on request of the Congress (Havens, 1992, pp. 208–9). Since the mid-1980s the GAO has shifted from mostly self-initiated work to congressionally driven work (O'Connell, 2007, pp. 26–7).
4. This list is not comprehensive. Without further empirical research it is not possible to gain insight into the way this factor impacts upon the selection process.
5. 'Overall, we believe the process is less than 15 per cent process, and is still more than 85 per cent judgement' (Office of the Auditor-General of Canada, 2002, 6); 'The

- identification of risk is not an exact science and there is no mechanistic formula for assessing whether something is at risk.' (National Audit Office, 1997, 7).
6. The description hereafter is based on the federal level within the Belgian Court of Audit.
 7. The Algemene Rekenkamer is, to our knowledge, the only SAI that makes 'audit synthesis', e.g. in the (meta-) audit 'Tussen beleid en uitvoering' ('Between Policy and Implementation'), 2003 (this report dates from *before* the new strategy), the synthesis (which was combined with a follow-up of previous recommendations) was based on 29 individual audits. This is a very interesting method to distil trends from individual audits, although there are also limitations: blind spots in the original audits will be blind spots in the audit synthesis; since audits are not randomly chosen, nobody can say if patterns in the synthesis reflect patterns in reality.
 8. Raaum & Morgan seem to assume that the auditable unit is also the unit of audit. This is not necessarily the case: after analysing risks by auditable unit, it is possible to reorganize this information to come to another scope.
 9. Some limitations can be overcome; others may turn out to be an inherent limitation to this technique.

REFERENCES

- Algemene Rekenkamer (2004a), *Onderzoek naar de aansluiting tussen beleid en uitvoering. Betekenis van de nieuwe strategie voor het onderzoeksproces*, The Hague: Algemene Rekenkamer.
- Algemene Rekenkamer (2004b), *Presteren en functioneren van het openbaar bestuur. Strategie Algemene Rekenkamer 2004–2009*, The Hague: Algemene Rekenkamer.
- Algemene Rekenkamer (2008), *Werkprogramma*, The Hague: Algemene Rekenkamer.
- Algemene Rekenkamer (2010), *Slagvaardig en transparant. Presteren en functioneren van het openbaar bestuur. Strategie Algemene Rekenkamer 2010–2015*, The Hague: Algemene Rekenkamer.
- Azuma, N. (2004), 'Performance measurement of Supreme Audit Institutions', *Government Auditing Review*, March, 65–94.
- Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale and Burt Perrin (eds) (2007), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Brunswick: Transaction Publishers.
- Bovens, M., P. 't Hart, and S. Kuipers (2006), 'The politics of policy evaluation', in Moran, M., M. Rein and R.E. Goodin (eds), *The Oxford Handbook of Public Policy*, Oxford: Oxford University Press, pp. 319–35.
- Chelimsky, Eleanor (2007), 'On fulfilling the accountability purpose in audit and evaluation: integrating politics and public involvement', in United Nations, *Auditing for Social Change: A Strategy for Citizen Engagement in Public Sector Accountability*, Department of Economic and Social Affairs, New York, pp. 115–42.
- Danish Audit Office (2008) 'Mission, vision 2012 and performance targets 2010', available at www.rigsrevisionen.dk/composite-489.htm.
- FrameWorks Institute (2002), *Framing Public Issues*.
- Government Accountability Office (2000), *Determining Performance and Accountability Challenges and High Risks*, Washington DC.
- Government Accountability Office (2003), *Transformation, Challenges, and Opportunities*, Statement of David M. Walker, Comptroller-General of the

- United States. Testimony Before the Senate Committee on Governmental Affairs, 16 September.
- Government Accountability Office (2004), *Strategic Plan 2004–09*, Washington DC.
- Government Accountability Office (2005), *21st Century Challenges. Re-examining the Base of the Federal Government*, Washington DC.
- Government Accountability Office (2007), *High Risk Series*, An update, Washington DC.
- Government Accountability Office (2009a), *Performance and Accountability Report 2009*, Washington DC.
- Government Accountability Office (2009b), *High Risk Series*, 99.
- Havens, H. (1992), 'The evolution of the general accounting office: from voucher audit to program evaluations', in Weiss, Carol (ed.), *Organizations for Policy Analysis. Helping Government Think*.
- House, E. (2006), 'Democracy and evaluation', *Evaluation*, **12** (1), 119–27.
- House, Ernest and Kenneth Howe (1999), *Values in Evaluation and Social Research*, Thousand Oaks: Sage.
- INTOSAI (2001), *Independence of Supreme Audit Institutions*, INTOSAI.
- INTOSAI (2004), *Implementation Guidelines for Performance Auditing. Standards and Guidelines for Performance Auditing based on INTOSAI's Auditing Standards and Practical Experience*, INTOSAI.
- Lonsdale, Jeremy (1999), 'Impacts', in Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilkka Summa and Marit Waerness, *Performance or Compliance? Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- Lonsdale, J. (2008), 'Balancing independence and responsiveness', *Evaluation*, **14** (2), 227–48.
- Lonsdale, Jeremy (2011), 'Using their discretion: how state audit institutions determine which performance audits to undertake', in Eliadis, Pearl, Jan-Eric Furubo and Steve Jacob (eds), *Evaluation: Seeking Truth or Power*, New Brunswick, N.J.: Transaction Publishers.
- Mazur, J. (1994), 'The SAIs right to select audit topics', *International Journal of Government Auditing*, April, 13–15.
- Murray, R. (2002), 'Citizens' control of evaluations', *Evaluation*, **8** (1), 81–100.
- National Audit Office (1997), *Picking the Winners: A Guide to VFM Study Selection*, London: National Audit Office.
- National Audit Office (2000), *Modernising Government, How the NAO are Responding*, London: National Audit Office.
- National Audit Office (2005), *State Audit in the European Union*, London: National Audit Office.
- National Audit Office (2006), *Corporate Plan 2006*, London: National Audit Office.
- National Audit Office (2008), *VFM Handbook. A Guide to Quality in VFM Work*, London: National Audit Office.
- National Audit Office (2009), *Annual Report*, London: National Audit Office.
- O'Connell, A.J. (2007), *Auditing Politics or Political Auditing?*, UC Berkeley Public Law Research Paper No. 964656, University of California, Berkeley.
- Office of the Auditor-General of Canada (2002), *One-Pass Planning at the Office of the Auditor-General of Canada*, International Seminar on Effective Performance Auditing, 30 and 31 May.

- Office of the Auditor-General of Canada (2004), *Performance Audit Manual*, Ottawa: Office of the Auditor-General.
- Patton, M.Q. (1997), *Utilization-Focused Evaluation: The New Century Text*, Thousand Oaks, California: Sage.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilka Summa, and Marit Waerness (1999), *Performance or Compliance? Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- Raam, R. and S. Morgan (2001), *Performance Auditing: A Measurement Approach*, The Institute of Internal Auditors.
- Rekenhof (2004), *Controleren, evalueren, informeren – Contrôler, évaluer, informer. Missieverklaring – Déclaration de mission*, Brussels: Rekenhof.
- Rekenhof (2006a), *Monitoring en risicoanalyse voor de selectie van auditthemas. Federale Sector*, Brussels: Rekenhof.
- Rekenhof (2006b), *Strategisch plan 2006–2009, 7 prioriteiten voor de federale sector*, Brussels: Rekenhof.
- Sigma (2002), *Relations between Supreme Audit Institutions and Parliamentary Committees*, Sigma Paper Series, no. 33.
- Sterck, M., V. Conings, V. Put and G. Bouckaert (2006), *De praktijk van performance audit: een drielandenstudie*, Steunpunt Bestuurlijke Organisatie.
- Turksema, R. and P. Van der Knaap (2006), 'Selectie en planning van doelmatigheidsonderzoek bij de Algemene Rekenkamer', *Vlaams Tijdschrift voor Overheidsmanagement*, nr. 4, 11–18.
- Van der Meer, F.B. (1999). 'Evaluation and the social construction of impacts', *Evaluation*, 5 (4), 387–406.
- Van der Meer, F.B. and J. Edelenbos (2006), 'Evaluation in multi-actor policy processes: accountability, learning and cooperation', *Evaluation*, 12 (2), 201–18.
- Weiss, Carol (1987), 'Where politics and evaluation research meet', in Palumbo, D.J. (ed.), *The Politics of Program Evaluation*, Sage Publications, 47–70.
- Weiss, Carol (1998), *Evaluation*, New Jersey: Prentice-Hall.
- Wilkins, Peter and Jeremy Lonsdale (2007). 'Public sector auditing for accountability: new directions, new tricks?', in Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale and Burt Perrin (eds), *Making Accountability Work. Dilemmas for Evaluation and for Audit*, New Brunswick: Transaction Publishers.

4. Norms used: some strategic considerations from The Netherlands and the UK

Vital Put

INTRODUCTION

What does the word ‘performance’ in performance audit mean? According to Pollitt (1999, 2003) performance auditors draw on notions of ‘good management’, yet the source and nature of these models of ‘good management practice’ tend to remain opaque (Pollitt et al., 1999, 204–5). This issue was identified by Pollitt and colleagues as meriting further research, and the object of this chapter is to examine further what Supreme Audit Institutions (SAIs) consider as ‘good performance’.

This chapter is based on detailed research designed to describe, characterize and explain the norms used in performance audits of SAIs. It is mainly based on a content analysis of audit reports of the Dutch Algemene Rekenkamer and the National Audit Office (NAO) in the United Kingdom. Both SAIs describe the norms they use in guidance or brochures. In the Algemene Rekenkamer’s document (Algemene Rekenkamer, 2005) the most important norms are described at a high level, whilst the norms used by the NAO are explained in several good-practice guides and in the section ‘Guidance & Good Practice’ on their website. However, we prefer here to make use of an analysis of audit reports as this provides unobtrusive evidence and captures the norms that were actually applied.

The performance auditing manual of INTOSAI defines norms (‘audit criteria’) as:

reasonable and attainable standards of performance against which the economy, efficiency, and effectiveness of activities can be assessed. They reflect a normative (i.e. ideal) model for the subject matter under review. They represent best or good practice, a reasonable and informed person’s expectation of ‘what should be.’ When criteria are compared with what actually exists, audit findings are generated. (INTOSAI, 2004, 103)

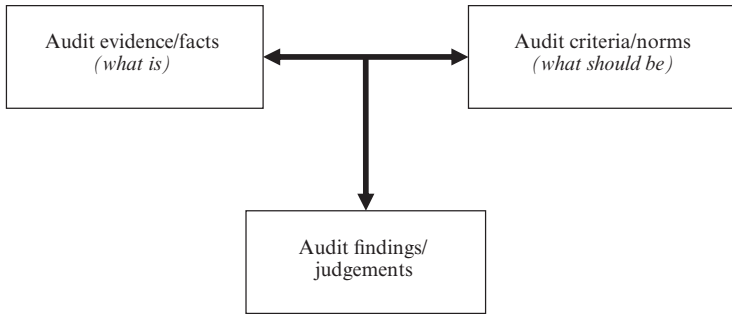


Figure 4.1 Facts, norms, judgements

According to these international standards, in performance audits an assessment is made of the effectiveness and efficiency of public organizations. Since assessments are based on a comparison of facts ('what is') against norms ('what should be'), norms are needed to evaluate performance. See Figure 4.1.

We start by considering why it is important for an SAI to be explicit about the norms it uses.¹ Firstly, at the level of individual audits, the auditee has the right to know on the basis of what standards his or her organization will be assessed (Pollitt, 1999, 102). Without explicit norms it is not possible to discuss (and possibly agree) these criteria with the organization. Secondly, without them, auditors will be uncertain as to what they are looking for and what constitutes effective performance, and they will not be able to discuss their ideas with auditees. In addition, a clear understanding of the norms in use is also necessary if we want to go beyond individual audits and assess the effectiveness of an SAI. This is especially important since the available scientific evidence about the relationship between the existence of certain 'good management practices' and the effectiveness and efficiency of policies and organizations is not conclusive.² Furthermore, the manuals of SAIs indicate that they are aware of these uncertainties.³ Finally, understanding the explicit norms also allows comparisons with those used by other scrutinizing bodies (for example internal auditors, evaluators, inspectorates), which will allow us to assess, amongst other things, if they are mutually consistent or if, by contrast, auditees are faced with incompatible requirements.

The development of norms is part of the audit planning process, which encompasses the identification of a topic (Chapter 3) and selection of methods (Chapter 5). During the preliminary study phase of an audit the norms will be elaborated and there will be a dialogue with the auditee about these norms. During the main audit process these norms will be

compared with the facts, in order to come to judgements. Often – but not always – the norms will be spelled out in the final audit report, sometimes in a general way, at others in more detail.⁴

To date, the norm-setting behaviour of SAIs has barely received attention from the scientific community (although see Bemelmans-Videc, 1998; Bemelmans-Videc & Fenger, 1999; Bemelmans-Videc, 2000). One writer who has examined the subject – Bemelmans-Videc – found four central norms in the performance audit reports of the Dutch Algemene Rekenkamer. These were that there should be insight into the effectiveness and efficiency of administration and policy; there should be a clear, well-considered and substantiated policy, that should be known to all relevant actors; there should be written reports (so that control is possible), and there should be accountability of ministers to parliament.

Apart from this, the subject has received only limited attention (for example, Pollitt et al., 1999; Radcliffe, 1999; Schwartz, 1999). Nevertheless, what evidence there is teaches us that the norms applied by SAIs, for the most part, concern the *internal functioning* of public organizations (for example management practices), and only to a much lesser degree the results of policies. They are often procedural (systemic) in nature, rather than substantive; are mostly qualitative; are very often implicit; are formulated unilaterally by SAIs, albeit they anticipate what is acceptable for the auditee; and they often come from official sources.

NORMS USED BY THE ALGEMENE REKENKAMER AND THE NATIONAL AUDIT OFFICE

In the next section, we describe the norms used by the Algemene Rekenkamer in The Netherlands and the National Audit Office in the United Kingdom. We use this factual material to formulate three, inter-related, strategic positions. Based on content analysis of audit reports of the two audit bodies, five categories of norms were found (Put, 2006a; Put 2006b). These were:

1. ‘good management’ practices (for example strategic planning, performance measurement, human resources management, information technology);
2. ‘good policy’ (for example the quality of policy objectives, evidence based policy);
3. ‘output/service delivery’ (e.g. quality, timelines);
4. ‘effectiveness’;⁵ and
5. performance accountability of the government towards the parliament.⁶

Table 4.1 Number of reports that address certain types of norms

	ARK (n=50)	NAO (n=60)
1. Norms regarding 'good management'	50%	92%
2. Norms regarding 'good policy'	88%	38%
3. Norms regarding 'output/service delivery'	32%	37%
4. Norms regarding 'effectiveness'	6%	3,3%
5. Norms regarding performance accountability of ministers towards the parliament	24%	-(¹)

Note: ¹ This category is about auditing the relevance and reliability of the accountability information that flows from ministers towards parliament. The publication of performance audit reports is, of course, an information flow in its own right, but this is not intended with this category. Since the NAO does not audit (the way) ministers (keep their parliament informed), this category is not relevant for the NAO.

Table 4.1 gives a quantitative overview of these categories.

As Table 4.1 shows, in a majority of reports, both SAIs look 'inside' the government machine, although at different things: the NAO, in particular, looks at management practices (in 92 per cent of its reports) and the Algemene Rekenkamer is strongly focused on policy practices (88 per cent of its reports). In almost a third of the reports of both SAIs, the output of public organizations is assessed (for example, the number of inspections to be carried out, timeliness for delivery of a licence, customer satisfaction etc). Effectiveness is seldom examined even though both SAIs formally state that they investigate it. In reality the effectiveness norm is barely applied (see also Schwartz, 1999, for similar conclusions), although importantly SAIs *do* address outputs and service delivery.

We now offer a more detailed description of these five categories of norms. The norms regarding good management and good policy (1 and 2) are considered as conditional norms, that is they contribute to good results, such as output and effectiveness (3 and 4).

The norms applied in performance audits by the Algemene Rekenkamer can be summarized as follows:

If there is a 'good policy', that is if ministers:

- formulate a clear, evidence-based, feasible and consistent policy (with clear and measurable objectives, a clear relationship between objectives – instruments, budget, timing) and evaluate the implementation and results of this policy periodically;

- run the public sector as an interconnected chain, a whole system – with clear responsibilities for each organization and clear co-ordinating arrangements;
- elaborate clear, consistent, complete, enforceable legislations, preceded by an ex-ante evaluation;
- supervise tightly their departments and agencies, that is, guide these organizations with a clear policy vision, monitor implementation and results of this policy; and adjust where necessary;

And if there is ‘good management’, that is if public managers:

- manage their organization with a substantiated, result-oriented planning and control cycle, aligned with the policy of the minister;
- establish a clear organizational structure, with a clear division of tasks and the necessary co-ordination mechanisms;
- ensure that cooperation with other organizations is arranged well;
- ensure that the primary processes within the organization are based on formal guidance and are documented;
- address IT-security in a systematic way (for example plans, risk assessments, and so on)

Then this will lead to ‘good results’, that is:

- sufficient services are delivered;
- services reach the intended target group;
- services are delivered on time;
- these outputs lead to the achievement of the official goals of the policy.

Furthermore: the responsible minister should inform parliament – in a complete, correct, timely, relevant way – about the implementation, results and costs of his policy.

For the NAO the norms can be summarized as follows.

If there is ‘good management’, that is, if public managers:

- manage their organization with a result-oriented, cost-conscious and customer-oriented strategy, established in consultation with staff and customers; subsequently monitor the implementation of this strategy with the help of a performance measurement system, and make adjustments if necessary;

- operate to a customer-oriented approach, i.e. consult with customers, obtain insight into their needs, handle and analyse complaints carefully and learn from it, communicate via user-friendly forms and websites;
- manage risks (i.e. anticipate implementation problems and provide for a 'plan B' in case things go wrong);
- learn from their own successes and failures, and learn from the experience of others (knowledge management);
- align their human resources management with their strategy, ensuring that there are competent staff;
- stimulate an organizational culture which supports risk taking, learning and joint working;
- align their procurement policy with their strategy, underpin procurement decisions with cost–benefit analysis and consultation with the concerned staff members, apply principles of risk management and project management for major acquisitions;
- tune their information technology (IT) to their strategy, underpin decisions about IT with good business cases, consult with users, apply principles of risk management and project management for IT projects, learn from previous projects;
- manage supplier relationships well;
- take care that operating processes are based on good practices, and establish quality assurance systems to safeguard the quality of these processes;
- provide a clear organizational structure;
- create good cooperation arrangements with other organizations;

Then this will lead to 'good results', that is:

- services are delivered to an appropriate standard, both from the customer's viewpoint (customers are satisfied) as well as that of the provider;
- services are delivered on time;
- services are delivered efficiently – in comparison with the past, with efficiency targets, with other organizations (benchmarking);
- there are sufficient services;
- these services will reach the intended target group and no groups of people are excluded.

These qualitative descriptions, in combination with the numbers in Table 4.1 above, show that the Algemene Rekenkamer strongly focused at the policy level (policy theory, legislation, oversight by ministers, and so on).

This is the responsibility of ministers and consequently the Algemene Rekenkamer scrutinizes in particular the activities of ministers. To a lesser degree it also assesses the managerial level. As such, it can be said that for the Algemene Rekenkamer the main norm underlying performance audit work is the democratic accountability of ministers about their policy. In contrast, the NAO exclusively examines the management level (strategic planning, implementation, performance measurement, and so on), as well as the relationship between government and customers (service quality). For the NAO, customer-oriented and efficient and effective management is the main norm. This is the responsibility of public servants and thus the NAO focuses primarily on their activities.⁷ The same picture emerges from the recommendations both SAIs formulate in their reports. The Algemene Rekenkamer mainly addresses its recommendations to ministers; the NAO addresses its recommendations solely to departments or agencies. Furthermore, the Algemene Rekenkamer employs a so-called 'contradictory debate' with the minister, who is held to account when the audit report is discussed in parliament. The NAO has a process known as 'clearance' (Sharma, 2007), firstly at working level and then with the leading officials of the department or agency to confirm the facts and their presentation. The most senior official, designated formally as the Accounting Officer, then has to give an account in the hearings of the Public Accounts Committee (PAC) in parliament. Thus, in contrast to the Algemene Rekenkamer, the NAO only deals with officials, not with politicians, and it is senior officials who are questioned by the parliamentary committee, not ministers.

In those reports where the sources of the used norms are mentioned, the Algemene Rekenkamer almost always refers to official sources of norms (for example, goals in policy documents, legislation or strategic plans; good practice in internal guidance of departments; handbooks of the Treasury). The NAO also refers to official sources (such as targets in public service agreements, guidance of the Cabinet Office), but, in addition to this, other sources are used. These include the results of surveys to identify good practices; comparisons (with past performance, similar organizations), customer satisfaction, professional literature, industry standards (such as private sector call centres). For example, in the NAO report 'Making a Difference. Performance of Maintained Secondary Schools in England' (2003), the auditors drew on the literature on school effectiveness and on focus groups (of teachers and head teachers) to identify good practices against which to compare performance. Recent guidance within the NAO (National Audit Office, 2009) warns to be cautious to adopt the auditees' own targets as yardsticks, and to compare these targets with those of similar organizations and with previous performance levels.

Most performance audit reports fail to explain why poor performance occurred, and do not contain explanatory information (in about 20 per cent of the reports of both SAIs attention was paid to explanatory factors). For the Algemene Rekenkamer this confirms the findings of previous research (De Vries, 2000, 237). If explanations are given, they are almost always very brief. There are several possible reasons for this (some of them will be discussed further in this chapter). In particular:

- a number of explanatory factors may relate to policy factors which fall outside the mandate of an SAI;
- demonstrating causal relationships is very difficult;
- the nature of the explanation depends on the auditor's assumptions on what causes good and bad performance;
- audits would become very time-consuming if auditors have to look for the deeper causes of every dysfunction they find;
- some auditors consider their job done after they have made an assessment about the performance of the auditee and they consider it the task of the auditee to look for explanations for under-performance.

Nevertheless, we found some excellent examples of audits that discussed causes of performance problems. For example, in the NAO report 'Tackling Pensioner Poverty: Encouraging Take-up of Entitlements' (2002), a lot of attention was paid to factors that explained why there was a take-up problem, whilst in the Algemene Rekenkamer report on Information Communication Technology (ICT) within the police force (ICT bij de politie, 2003) there was a substantial discussion of the causes of ICT problems. In recent reports and manuals the tide seems to be turning in both offices. For example, in its strategy for 2004–09,⁸ the Algemene Rekenkamer puts more emphasis on explaining the gap between policy and implementation, and in its most recent performance audit manual⁹ explicitly addresses the need to gain insight into the causes of poor performance, giving examples of factors that could explain unsatisfactory outputs, effects or efficiency (such as a mismatch between goals, resources and time; poor legislation; a wrong policy theory; conflicting policy goals; insufficient support from stakeholders).

It is also worth mentioning a recent study on this subject. Siemiatycki (2009) made a comparison of studies by academics and audit reports of government auditors on cost overruns during the delivery of transportation infrastructure projects. He found that there were sharp divergences between these two groups regarding the kind of explanations used to explain cost overruns. Auditors tended to prioritize technical and managerial explanations (such as inadequate forecasting techniques, scope

changes, poor project reporting), whilst the academic literature prioritized political, economic, and psychological explanations, (for instance, deliberately underestimating project costs during the planning stage to build support; and optimism bias).

A MANAGERIALIST–CONSUMERIST VERSUS A POLITICAL APPROACH TO PERFORMANCE AUDITING

Streim (Streim, 1994) states that the demand for auditing arises because of the existence of agency problems;¹⁰ without such problems, auditing is superfluous. In Figure 4.2 the public audit system in a parliamentary democracy is conceived as a chain of principal–agent relations.¹¹ In each relation shown in the diagram, audit could be used to solve agency problems.

We can distinguish four actors in Figure 4.2: the people (in a dual role, as citizen (voter–taxpayer), and as a client and user of services), the parliament (elected representatives), the government (ministers) and the administration (departments, agencies). A principal–agent relation exists between the citizens and the parliament. Citizens elect parliamentarians and hold them accountable through the electoral process. It is rather unusual to have auditors to solve this agency problem directly; instead this part of the chain is covered by the mass media, political parties, advocacy groups,¹² although auditors may play some type of role if their work is used by any of these groups. Parliamentarians in turn approve government spending (the budget) and hold the government accountable for its performance in spending the money voted. A principal–agent relation exists between the parliament and the government (ministers). Both relations are political in nature.

Next there is a principal–agent relation between the government (ministers) and the administration. The role of officials in the administration is to implement policy decided by ministers. And finally, people – as clients of public services – can also hold service providers directly accountable for performance (for example, through complaints processes, by taking a matter to court or by choosing another service provider where there is competition). These relations are managerial or legalistic in nature.

Each relationship shown in Figure 4.2 requires different types of audit products. Members of Parliament will be interested, in particular, in what government does, and in the results of government policy. This knowledge can be obtained from, amongst other sources, performance audits. Ministers will be more interested in the implementation of their policy by

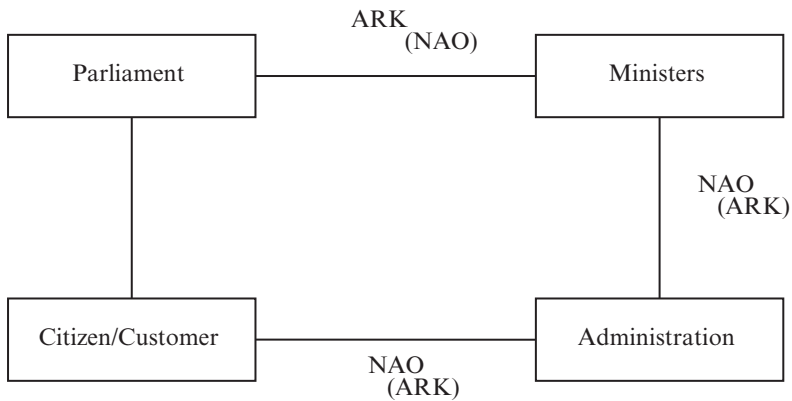


Figure 4.2 The public audit system as a chain of principal–agent relationships

the administration, and thus might be more influenced by evaluations and performance data provided by officials. Clients will be interested in service delivery, perhaps via published reports of performance.

Where can we pinpoint the Algemene Rekenkamer and the NAO in this diagram? As demonstrated earlier, the Algemene Rekenkamer chiefly focuses on the policy level and on holding ministers accountable to the parliament. It formulates norms for its audits relating to how ministers should make policy, on how ministers should give their accounts to parliament, on how ministers should supervise their administration and, to a lesser degree, on how public managers must manage their organizations and deliver services. As such, we have situated the Algemene Rekenkamer primarily in the principal–agent relation between the parliament and ministers, and only secondly in the relations between government/administration and administration/client. The NAO, on the other hand, examines the management level and the relation with customers. Therefore, we positioned the NAO primarily in the relationship between government and administration, and between administration and client, and finally in the relationship between ministers and parliament. The positioning of the Algemene Rekenkamer fits in with its strong focus on accountability for policy delivery, that of the NAO is consistent with its aim of contributing to the improvement of the performance of the public sector.¹³

This diagram can also be used to diagnose ‘strategic audit failures’, or mismatches between the position of the auditor in the principal–agent chain and the type of audit product offered by the auditor. For example, there would be a mismatch if an SAI strives to support the accountability

of ministers towards parliament, but delivers audit reports which only report on performance in terms of processes and do not offer information on results.

EXPOSING VERSUS EXPLAINING PERFORMANCE PROBLEMS

As mentioned earlier, most performance audits do not contain much information that explains why things went wrong, as opposed to what went wrong. This raises the question of the relevance of an explanatory analysis in a performance audit report. Majone offers some interesting lines of thought in this respect. He states (Majone, 1989, 177) that the choice of the type of evaluation should depend on the measurability of the outcome and on the predictability of the production function (the relationship between process and outcome) (Table 4.2, below).

Using this thinking, if there is complete knowledge of the relationship between process (for example certain management practices) and outcomes, an audit can be restricted to processes (quadrants I and II in the table) because there will be certainty that the presence of certain processes goes together with presence of the desired outcomes. Even then, however, in order to make the audit findings more appealing and to demonstrate their relevance, it is appropriate to estimate the effects of the findings on service delivery and the impact on society. For example, auditors could show how weaknesses in staff planning lead to a lack of competent staff members and how this leads to weak service delivery, or how weaknesses in quality management systems indeed cause quality problems. In order to be useful, it is also necessary to explain why dysfunctional processes occur.

Table 4.2 Choice of type of evaluation

		Knowledge of process	
		Complete	Incomplete
Measurability of outcomes	High	I Evaluation by process or by outcome	III Evaluation by outcome
	Low	II Evaluation by process	IV Evaluation by input

Source: Majone, 1989.

Very often, however, public policymakers and managers work in a context with fuzzy means–end relationships (Mintzberg, 1999; Noordegraaf, 2002; Wholey, 1999). In that case it is not a valid approach to look inside the box, because good management practices will not guarantee that goals will be achieved. As has been observed elsewhere, ‘if the connection between good management practices and the “three Es” is less than sure, then performance auditors will need to go beyond the former in order directly to investigate the latter’ (Pollitt et al., 1999, 206). Or, as Majone puts it: ‘Unless a definite relationship between inputs and outputs or outcomes – a well-defined production function in the language of the economist – can be assumed, input variables are a poor proxy for what we are really interested in knowing, namely, how effective is a given program, or how good is a particular policy’ (Majone, 1989, 173).

As a result, in that case, auditors should move to quadrant III in the matrix above and provide a direct measure of outputs and outcomes. But a pure black-box approach is also not sufficient. Information about results, without information that explains shortfalls in the observed results, does not provide a basis for informed decisions. Therefore, it is essential to complete an output/outcome approach, with an explanatory analysis. As Weiss states: ‘More learning will come from specifying the conditions that are associated with better or poorer outcomes – conditions of program operations as well as conditions in the larger social context’ (Weiss, 1987, 63). In order to provide a valid explanation it is important to go beyond the common sense knowledge, that is available to the average auditor, and to proceed in a structured way, making use of the best available scientific evidence to explain dysfunctions in the politico-administrative system and/or (negative) evolutions in the nature, size and distribution of societal problems.

The remaining category, quadrant IV, is difficult to audit, since in this situation there are no measurable objectives and since the presence of certain processes does not guarantee that the objectives will be achieved. If we take these elements into consideration, the most appropriate approach can be pictured as in Table 4.3.

The line of reasoning above starts from the premise that an SAI wants to contribute to improving the performance of the public administration. This requires more ‘diagnostic’ performance audits in order to formulate relevant recommendations. This poses huge challenges for SAIs. For example:

- Some explanatory factors may relate to political factors that fall outside the legal mandate of an SAI (such as political compromises). For example, a recent NAO report which looked at the costs and

Table 4.3 Choice of the most valid type of performance audit

		Knowledge of process	
		Complete	Incomplete
Measurability of outcomes	High	Audit of processes, including: <ul style="list-style-type: none"> ● an estimation of the consequences of dysfunctional processes; 	Audit of results, followed by an analysis to explain shortfalls in results
	Low	<ul style="list-style-type: none"> ● an explanatory analysis to explain dysfunctions. 	Difficult to audit in a meaningful way

benefits of the restructuring of government bodies did not comment on the fact that several major reorganizations were done to provide large ministries for senior ministers.

- Explanatory information is less defensible because demonstrating causal relationships is always tricky. Jan-Eric Furubo’s chapter in this book has made it clear that because of the role performance audit often plays in accountability mechanisms, the expectations about the robustness of the evidence are very high.
- The nature of the explanation depends importantly on the auditor’s assumptions on what causes good and bad performance. For example, auditors who adhere to rational decision models will look for explanations in weaknesses in planning systems, unclear responsibilities, shortcomings in information systems, absence of adequate tools; auditors who prefer internal control models will explain weaknesses by an inadequate internal control system (for example, inaccurate and incomplete risk assessment, inadequate contingency plans etc); advocates of public choice theory will look at incentive structures (for example, no incentive to listen to customers because of a monopoly position; no link between performance and funding, and so on); proponents of human resources models will refer to skills, communication, organizational culture etc. Explaining, explicitly or implicitly, starts from a certain model about the mechanics of organizations, policy etc. Each model makes certain variables visible, but ignores others (Bolman and Deal, 2003; Bovens and ’t Hart, 1998; Harrison and Shirom, 1999; Morgan, 1992; Scott, 2003).

From an accountability point of view, explanatory information is less relevant. An SAI could take the strategic decision only to expose problems, without analysing the causes of these problems and without providing

solutions. In that case, performance audits can be done in a purely ‘checking’ mode: comparing facts with norms and coming to a judgement based on this comparison. If managers or politicians want to use such performance audits to improve corporate performance, they will need an additional audit or evaluation by another body to deliver more useful information.

AN OFFICIAL VERSUS PLURALIST APPROACH TO NORMS

We have seen that SAIs very often refer to official sources of norms. Frequently auditing concerns establishing whether an auditee has complied with official goal statements; in effect, what government has said it has set out to achieve. From an accountability point of view, official norms are very relevant, particularly policy objectives set by democratically elected politicians, who have to give account for what they promised to do. In a democracy all power belongs to the people, the people elect political representatives to make decisions for them, these representatives delegate their decision-making power to governments, and governments in turn delegate the power to take certain decisions to civil servants. Since official norms are rooted in a statement authorized somewhere in the democratic chain, they have a strong democratic legitimacy (Vedung, 2006). From this perspective, an SAI checks whether the auditee has complied with the auditee’s own definition of performance: policy and management are judged by their own norms (‘official’ norms).

Moreover, starting from official norms also has some practical advantages. For example:

- official norms are the most defensible ones since they come from the auditee, who will find it hard to contest them; and
- official sources are also the cheapest way to formulate norms, given that SAIs do not have the resources to establish a Research and Development division to monitor scientific research thoroughly; this is an important advantage.

If an SAI wants to contribute to a better functioning public sector, however, it is less clear that it can rely simply and solely on official norms. There are a number of reasons why this might be the case. For example:

- An SAI that wants to act as a pioneer in introducing new ideas about public management cannot wait until these ideas are absorbed in official sources of norms.

- It could be questionable to start from official policy objectives if they are unrealistic by being either too easy to achieve or too ambitious.
- Using official norms also entails the risk of undesirable side-effects, namely ‘gaming’: the auditee could formulate objectives that are easy to attain if he/she knows that these objectives will be the starting point for audits.
- Only assessing against official norms gives an impoverished image of reality and the perspectives of other stakeholders are also valuable (for example, norms formulated by customers, peer organizations). Thus, some scholars argue for a more pluralistic approach (Majone, 1989, 171; Chelimsky, 2007, 121), by including the values, expectations and needs of all stakeholders. According to these authors, this approach would give a richer picture of reality.
- In some circumstances there are no official norms and SAIs have no choice but to construct their own.

From this viewpoint an SAI has to go beyond official norms and, for example, use scientific knowledge, benchmarking with other countries and international sources of norms, and so on. In this way performance is defined by the SAI, something which requires expert knowledge. From an accountability point of view this option could be seen as a ‘mismatch’. If the audit office uses its own norms on performance, instead of official ones which are approved by the principal, then the auditor acts as a principal. Audit then gets confused with accountability, instead of being a system that is superimposed on an accountability relationship. As Glynn states: ‘External audit in its various forms is essentially an *ex post* process that is superimposed on an accountability relationship that involves at least two parties, the one who allocates responsibility and the one who undertakes to discharge that responsibility. . . . It is not or should not be a substitute for accountability’ (Glynn, 1996). So there is clearly a dilemma for SAIs.

CONCLUSION: PERFORMANCE AUDITING ‘FOR DEMOCRACY’ AND/OR ‘FOR IMPROVEMENT’?

In the paragraphs above we sketched three fundamental strategic options. Each time we made a link with the double mission of SAIs: informing parliament and encouraging improvement in public administration. The three strategic positions are not independent of each other; rather they match with an accountability mission or with a mission to improve public administration, as is shown in Table 4.4.

Most SAIs try to adhere to both positions. The accountability approach

Table 4.4 Auditing for accountability versus auditing for improvement

Performance auditing for accountability	Performance auditing for improvement
<p><i>Political approach</i> The SAI delivers information that is useful for accountability purposes (i.e. in the relation between citizens and parliament and/or – commonly – the relation between parliament and government).</p>	<p><i>Managerial–consumerist approach</i> The SAI delivers information that is useful to improve the functioning of the public administration (i.e. information about the functioning of the administration and/or information about the relation between clients and administration).</p>
<p><i>‘Revealing’ performance audits</i> Exposing problems, without analysing the causes of these problems and without providing solutions. Performance audits are done in a purely ‘checking’ mode.</p>	<p><i>‘Diagnostic’ performance audits</i> Exposing and explaining problems (in order to formulate recommendations to solve these problems).</p>
<p><i>‘Official’ norms</i> The SAI checks whether the auditee has complied with its own definition of performance, with its own ‘promises’.</p>	<p><i>The SAI formulates its own definition of what ‘performance’ should be</i> Next to official norms, the SAI makes use of professional literature, benchmarking, the views of other stakeholders, . . .</p>

towards performance auditing fits within the traditional mission of SAIs to assist their parliament in holding government accountable; it does not require extensive expert knowledge. The second option is much more challenging, it requires a more complex knowledge base (such as the ability to assess effectiveness, knowledge of organizational theory). In addition, this commitment to performance improvement – if taken to its logical conclusion of helping organizations – does not fit very well with the institutional position of SAIs (see also Barzelay, 1997; Morin, 2003; Saint-Martin, 2004). An accountability relationship is a triangular relationship (for example, parliament/principal, auditee/agent and auditor), premised on distrust, in which the auditee is an object of audit, and the focus is on accounting for the past. A supportive relationship on the other hand, is a one-to-one relationship, based on trust, in which the auditee is a subject that takes part in deciding what happens, and the focus is on problem solving for the future.

Even when an SAI emphasizes its improvement mission, it continues to

function in an accountability triangle: the mandate of SAIs often emphasizes accountability, and the knowledge base and culture of performance auditors are highly influenced by those of financial statements audit that is very focused on accountability (see: Barzelay, 1997; Morin, 2003; Saint-Martin, 2004). This creates tensions between competing norms. Mulgan (2001) states that values such as accountability and transparency may set up tensions with the more managerialist approach, especially on issues such as outsourcing (for example, with an over-use of ‘commercial-in-confidence’ disclaimers). Funnel (2003) states that ‘in the exaltation of the three E’s of economy, efficiency and effectiveness, tensions between the individual as citizen and the individual as consumer/customer can be forgotten easily.’ SAIs have to learn to deal with these ambivalences.

NOTES

1. However, not all performance audits have to be criteria-based. The INTOSAI guidelines for performance auditing distinguish two approaches in examining performance: a result-oriented, normative approach (which assesses whether pre-defined criteria/norms have been achieved) and a problem-oriented approach (which starts with a stated problem and verifies and analyses the causes of this problem, without reference to predefined audit criteria) (INTOSAI, 2004, 26–7).
2. For example, in their book on managerial effectiveness Bozeman and Straussman state that ‘the guidelines we present have little to do with science. Some are loosely connected to empirical research, but in most cases there is no firm research base, much less a scientific theory basis, for the guidelines’ (1990, 25). Hood has commented ‘It is now commonly recognized that the ‘science’ of public management – indeed management generally – is heavily rhetorical . . . Most what-to-do arguments in public management rely on circumstantial evidence and rhetorical power’ (Hood, 2000, 13). There is a relevance gap between management research and the practice of management (Starkey and Madan, 2001; Hodgkinson et al., 2001). Tranfield argues for the development of ‘evidence-based’ knowledge about management, with the help of systematic reviews of research, like in medical science (Tranfield et al., 2002).
3. For example, National Audit Office, VFM Guide, 1997, 37; and Algemene Rekenkamer, Handleiding Doelmatigheidsonderzoek, 1996, 15–16.
4. In our research we found that norms were (at least partly) spelled out in about half of the reports.
5. Effectiveness is defined as the causal relationship between the output delivered and the effects in society (‘policy effectiveness’), and not as the causal relationship between activities and output (‘managerial effectiveness’).
6. This norm is mentioned separately because a transparent, democratic accountability of ministers towards the parliament is a goal in itself for the AR, and not (just) a condition for effectiveness and efficiency.
7. Some reports of the NAO concern policymaking, but they focus on the role of civil servants in the policymaking process and good practice, for example, in the use of research in policymaking.
8. Algemene Rekenkamer (2004), *Presteren en functioneren van het openbaar bestuur. Strategie Algemene Rekenkamer 2004–09*, p. 28.
9. Algemene Rekenkamer (2006), *Manual Performance Audit*, p. 198 (see e.g. pp. 25–8, 113–15).

10. The agent not acting in the interest of the principal but in its own interest.
11. This diagram suffices for our purpose, but reality is somewhat more complex. For example, departments and agencies can utilize private actors (non profit organizations, firms) to actually provide public services, this extends the principal-agent chain; there are also principal-agent relationships within the administration (for example, most internal audit services are situated within the relationship between management and employees).
12. But SAIs' work can provide the basis for the mass media and others to comment.
13. In its mission statement the NAO strives both for enhancing accountability to parliament and helping audited bodies to improve their performance. There are several indications that the second part of its mission is more dominant (at least in its performance audit work). In its 2003 manual the NAO states that promoting beneficial change is the ultimate goal of VFM work (Value for Money Handbook. A guide for building quality into VFM examinations, 2003, 10); the performance measures used by the NAO to measure its own performance (saving money, implementing recommendations); and the way the NAO justifies its reason of existence (Pollitt & Summa, 1997) also relate to improvement objectives.

REFERENCES

- Algemene Rekenkamer (2005), *Essentialia van goed openbaar bestuur*, The Hague.
- Barzelay, M. (1997), 'Central audit institutions and performance auditing: A comparative analysis of organizational strategies in the OECD', *Governance: An International Journal of Policy, Administration and Institutions*, **10** (3), 235–60.
- Bemelmans-Videc, M.L. (1998), 'De Algemene Rekenkamer: controlenormen en – stijlen in een veranderende bestuurlijke context', in Hertogh, M.L.M., Huls, N.J.H. and Wilthagen, A.C.J.M. (eds), *Omgaan met de onderhandelende overheid. Rechtstaat, onderhandelend bestuur en controle*, Amsterdam University Press, pp. 89–117.
- Bemelmans-Videc, M.L. (2000), 'Over het beoordelen van "goed" bestuur', in Nelissen, N., Goverde, H., Van Gestel, N. (eds), *Bestuurlijk vermogen, Analyse en Beoordeling van Nieuwe Vormen van Besturen*, Uitgeverij Coutinho.
- Bemelmans-Videc, M.L. and H.J.M. Fenger (1999), 'Harmonizing competing rationalities in evaluating governance', *Knowledge, Technology & Policy*, **12** (2), 38–51.
- Bolman, Lee and Terrence Deal (2003), *Reframing Organizations. Artistry, Choice and Leadership*, Third Edition, San Francisco: Jossey-Bass.
- Bovens and 't Hart (1998), *Understanding Policy Fiascoes*, New Brunswick, Transaction.
- Bozeman, Barry and Jeffrey Straussman (1990), *Public Management Strategies. Guidelines for Managerial Effectiveness*, San Francisco: Jossey-Bass Publishers.
- Chelimsky, E. (2007), 'On fulfilling the accountability purpose in audit and evaluation: Integrating politics and public involvement', in *United Nations, Auditing for Social Change: A Strategy for Citizen Engagement in Public Sector Accountability*, Department of Economic and Social Affairs, New York, 2007, pp. 115–42.
- De Vries, G.J. (2000), *Beleidsdynamica als sociale constructie. Een onderzoek naar de doorwerking van beleidsevaluatie en beleidsadvisering*, Delft: Eburon.

- Funnel, W. (2003), 'Enduring fundamentals: Constitutional accountability and Auditors-General in the reluctant state', *Critical Perspectives on Accounting*, **14** (2), 107–32.
- Glynn, J.J. (1996), 'Performance auditing and performance improvement in government: Public sector management reform, changing accountabilities and the role of performance audit', in OECD, *Performance Auditing and Public Sector Modernization*, Paris, pp. 125–36.
- Harrison, M.I. and A. Shirom (1999), *Organizational Diagnosis and Assessment. Bridging Theory and Practice*, Thousand Oaks, California: Sage Publications.
- Hodgkinson, G.P., P. Herriot and N. Anderson (2001), 'Re-aligning the stakeholders in management research: Lessons from industrial, work and organizational psychology', *British Journal of Management*, **12**, S41–S48.
- Hood, Christopher (2000), *The Art of the State: Culture, Rhetoric and Public Management*, Oxford: Clarendon.
- INTOSAI (2004), *Implementation Guidelines for Performance Auditing. Standards and Guidelines for Performance Auditing Based on INTOSAI's Auditing Standards and Practical Experience*. INTOSAI (www.issai.org).
- Majone, Giandomenico (1989), *Evidence, Argument and Persuasion in the Policy Process*, Yale University Press.
- Mintzberg, Henry (1999), *Grandeur et décadence de la planification stratégique*, DUNOD.
- Morgan (1992), *Beelden van Organisaties*, Schiedam, Scriptum.
- Morin, D. (2003), 'Controllers or catalysts for change and improvement: Would the real value for money auditors please stand up?', *Managerial Auditing Journal*, **18** (1), 19–30.
- Mulgan, R. (2001), 'Auditors-General: Cuckoos in the managerialist nest?', *Australian Journal of Public Administration*, **60** (2), 24–34.
- National Audit Office (2009), *Coming to a Sound Conclusion, A Guide to Developing Strong Judgements in Value for Money Reports*, London: National Audit Office.
- Noordegraaf, M. (2002), 'Wat doen publieke managers? De realiteit achter de planning & control-retoriek', *Bestuurskunde*, Number 1, 2–12.
- Patton, Michael (1997), *Utilization Focused Evaluation*, Edition 3, Beverly Hills, California: Sage Publications.
- Pollitt, C. (2003), 'Performance audit in Western Europe: Trends and choices', *Critical Perspectives on Accounting*, **14**, 157–70.
- Pollitt, C. and H. Summa (1997), 'Reflexive watchdogs? How supreme audit institutions account for themselves', *Public Administration*, **75**, 313–36.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilkka Summa and Marit Waerness (1999), *Performance or Compliance? Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- Put, V. (2006a), *De bril waarmee auditors naar de werkelijkheid kijken. Over normen die rekenhoven gebruiken bij het beoordelen van de overheid*, die Keure, Brugge, 219 p.
- Put, V. (2006b), 'Through which lens do Supreme Audit Institutions assess the performance of the public sector?', Paper presented at the 4th International Conference on Accounting, Auditing and Management, The European Institute for Advanced Studies in Management (EIASM), Siena, 7–9 September.
- Radcliffe, V.S. (1999), 'Knowing efficiency: The enactment of efficiency in efficiency auditing', *Accounting, Organizations and Society*, **24**, 333–62.

- Saint-Martin, D. (2004), 'Managerialist advocate or "control freak"', The Janus faced Office of the Auditor General', *Canadian Public Administration*, 121–40.
- Schwartz, R. (1999), 'Coping with the effectiveness dilemma: Strategies adopted by state auditors', *International Review of the Administrative Sciences*, **65** (4), 511–26.
- Scott, W.R. (2003), *Organizations. Rational, Natural and Open Systems*, Prentice Hall.
- Sharma, N. (2007), 'Interactions and interrogations: Negotiating and performing value for money reports', *Financial Accountability and Management*, **23** (3), 289–311.
- Siemiatycki, M. (2009), 'Academics and auditors. Comparing perspectives on transportation project cost overruns', *Journal of Planning Education and Research*, **29**, 142–56.
- Starkey, K. and P. Madan (2001), 'Bridging the relevance gap: aligning stakeholders in the future of management research', *British Journal of Management*, **12**, S3–S26.
- Streim, H. (1994), 'Agency problems in the legal political system and Supreme Auditing Institutions', *European Journal of Law and Economics*, **3**, 177–91.
- Tranfield, D., D. Denyer and P. Smart (2002), *Developing an Evidence-based Approach to Management Knowledge using Systematic Review*, Advanced Management Research Centre, Cranfield School of Management.
- Vedung, E. (2006), 'Evaluation research', in Peters, B.G. and Pierre, J. (eds), *Handbook of Public Policy*, London: Sage, pp. 397–417.
- Weiss, C.H. (1987), 'Where politics and evaluation research meet', in Palumbo, D.J. (ed.), *The Politics of Program Evaluation*, Beverly Hills, California: Sage Publications, pp. 47–70.
- Wholey, J.S. (1999), 'Performance-based management. Responding to the challenges', *Public Productivity and Management Review*, **22** (3), 288–307.

5. The right tools for the job? Methods, choice and context

Jeremy Lonsdale

Audit is an evidence-based activity. As Flint (1988) puts it ‘In no case can an auditor express an opinion or make a report without making an investigation and obtaining evidence.’ Thus, audit work is directed at gathering the evidence to enable the auditor to form a judgement that can be reported; if ‘there is no evidence an audit is not possible’ (Flint, 1988). The methods used by auditors determine what evidence they uncover, what analysis they can undertake and ultimately what conclusions and recommendations they are able to produce. They are thus fundamental to the value of performance audit work. This chapter considers the choice of methods used by performance auditors at the National Audit Office (NAO) in the United Kingdom. It does so by looking back at developments over a 20-year period, during which time performance audit (known as value for money audit in the United Kingdom) has changed considerably.

The starting point for this chapter is the view that a key determinant in the choice of methods in activities such as audit or evaluation is the context in which the work is done. The introductory material for the 2009 American Evaluation Association conference stated:

context [refers] to the setting (time and place) and broader environment in which the focus of the evaluation (evaluand) is located. Context can also refer to the historical context of the problem or phenomenon that the program or policy targets as well as the policy and decision-making context enveloping the evaluation. Context has multiple levels and is dynamic, changing over time. Increasingly, we are aware of the need to shape our methods and overall approach to the context. Each of the dimensions within the context of an evaluation influences the approaches and methods that are possible, appropriate, and likely to produce actionable evidence. (American Evaluation Association, 2009)

This perspective suggests there may be a complex set of processes at play, with the context in which audit work is undertaken influencing the

methods used, and auditors adopting particular methods to take account of the context. And these processes are subject to constant change. We start by considering the context in which the NAO has operated.

VALUE FOR MONEY AUDIT: CONTEXT AND TRENDS

The National Audit Office (NAO) in the United Kingdom has produced VFM reports in their current format since 1984 and in recent years has published 60 each year. Although individually designed, they usually share a number of common characteristics. They are discrete reports on aspects of the use of public money; the work has generally been carried out by a blend of in-house staff and external suppliers, and involves both quantitative and qualitative methods; audits focus on three or four sub-issues, generally reflected in a similar number of chapters in the final report; considerable attention is given to presentation, with a mix of graphics and case examples; and they are written for the non-specialist reader, with an assumption that they will form the basis of a parliamentary hearing at which senior officials will be questioned. Reports are ‘agreed’ in the sense that the facts and their presentation are discussed with the organisation examined, and their comments sought and taken into account in finalising the report (Sharma, 2007).

The NAO works within formal accountability structures. As was mentioned in Chapter 1, it reports to the House of Commons Public Accounts Committee (PAC), which produces its own reports on the basis of the NAO’s work and the oral evidence provided by senior civil servants at the PAC hearing. Around 50 VFM reports a year form the basis for hearings of the Committee. The government responds formally in writing to the Committee’s recommendations. This response either sets out the changes made to meet PAC recommendations or (less often) explains why they are not accepted. Comments and observations made in NAO reports are accorded significance, and are seen in parliament as matters to which senior administrators should pay attention.

VFM audits have been undertaken within increasingly demanding parameters. NAO produces more reports now than in the late 1990s (up to 60 each year from 45 as a result of growing demands from the PAC), and study teams have less time (averaging 9 months from approval to publication, with strong internal pressure from 2009 onwards to reduce this considerably for many projects) to develop a study, secure agreement to doing it, collect evidence, produce their reports and discuss them with the departments concerned. The presentation of NAO reports has received

attention in recent years, with greater pressure to produce shorter main reports (now 10,000 words maximum), and make them more readable for the non-specialist. Media attention has grown considerably, and many reports receive prominent coverage in the national press or on television, radio and the internet. All these are important contextual factors within which decisions about methods are made.

There is also pressure from different quarters which affects the conduct of the work. Lonsdale and Mayne's (2005) comparative analysis of quality arrangements in performance audit in a number of countries indicated that quality means different things to different audiences, be they politicians, officials within audited bodies, the media, subject specialists or citizens. What is deemed to be appropriate quality by one group may not necessarily meet the expectations of others. Thus, for example, civil servants might value a detailed analysis of a problem and prefer to see more critical comments about past performance set in a wider and more forward-looking and positive context, whilst journalists, looking at a report for the purposes of writing an article on government spending, might expect greater brevity and a more critical tone. Sector specialists might be interested in a report with detailed appendices.

The credibility of the NAO's VFM work is heavily dependent on recipients having confidence in the rigour of the analysis. The PAC expects the NAO to get to the bottom of particular issues, and the audit work is more likely to be useful as a basis for holding departments accountable if it is soundly-based and convincing. Equally, as external auditors, the NAO can be of most assistance to those organisations it audits only if the recommendations are based on rigorous work. The NAO is thus vulnerable to criticism from external commentators if it does not have robust and compelling messages based on thorough and convincing audit and research.

A historical perspective is also relevant. Up until the 1970s, VFM work consisted mainly of the examination of large schemes or projects and a wide range of transactions. Studies were narrowly focused and few topics arose from an analysis of risk or an evaluation of departmental systems. It was often difficult to distinguish between financial audit and VFM work, not least because the methods used were similar. Many staff regarded their work as 'point hunting' (mostly involving examination of documents) to identify individual failures, and tended to focus on 'the more obvious points of economy, while deeper questions of efficiency and effectiveness were less frequently pursued.' (Exchequer and Audit Department, 1981, 5.)

The National Audit Act 1983 gave a new impetus to how VFM work was carried out. It allowed reports to be published in a more timely manner, provided for wider focused audit (the legislation explicitly allowed a focus on economy, efficiency and effectiveness for the first time)

and, by allowing the Comptroller and Auditor General to present reports on individual topics rather than issues buried away with the accounts, encouraged the development of a new form of audit. Thereafter, successive Comptroller and Auditors General argued their work was concerned, less with individual failure and cost saving, and more about identifying better ways of doing things (Lonsdale, 2000). Studies were designed to 'add value', with NAO reporting at times on innovations in government and the implementation of new policies at an early enough stage to hold lessons for later stages. On issues like administrative reorganisations, IT contracts and programmes, private finance contracts or public-private partnerships, NAO reports in one particular area often have wider application to other government bodies tackling similar issues. Such work and the changing aspirations associated with it required different approaches to evidence gathering and analysis.

STUDY DESIGN AND CHOICE OF METHODS

The selection of methods by auditors should be considered within the context of how an overall study is developed. In general, NAO teams design their studies having secured approval for the topic from the Comptroller and Auditor General (C&AG), as head of the NAO, in whose name all the work is undertaken. There are a number of different planning considerations, including:

- the scoping of the subject and key questions,
- the creation of project management templates and plans,
- establishment of relations with the audited body to agree practical arrangements such as access to staff and visit schedules,
- consultation with internal and external stakeholders and experts,
- the exposure of plans to internal peer review and their refinement in the light of comments,
- the securing of senior management approval to the plans, budget and timetable,
- the identification of possible areas for use of external contractors, and
- the selection of appropriate methods to generate evidence (National Audit Office, 2008a).

Such a mix of considerations supports the view of some that developing a VFM study is a very practical exercise. Pollitt and Summa (1996) argued that auditors appeared to 'plan' their audits, whereas evaluators

‘designed’ their studies. In their view, it was rare for audit examinations to involve such self-consciousness about the way different methodological components worked together. Instead, ‘more commonly, audit teams will plan their audits in a fairly pragmatic fashion, selecting on a more-or-less ad hoc basis whatever tools and methods seem potentially useful for shedding light on the chosen program or institution.’ Linked to this, Lonsdale (2000) noted ‘although most auditors have had a professional accountancy training, they appear not to be tied to a particular set of concepts or approaches. Many choices are based on experiences learnt on the job, and on seeing evidence of what has been successfully carried out by others.’

More recent evidence from research with teams suggests designing a study is also a process of negotiation between team members. Junior members of teams prepare early versions of study plans, which are discussed with their managers, who may add in or remove particular methods, depending on their view of the needs of the study. Some of this may involve anticipating the attitudes (positive and negative) of more senior staff, based on past experience or on early discussions on the focus of the study. In a survey of 32 NAO VFM managers (the majority of those responsible for leading VFM studies), undertaken by the author, managers themselves and their teams were considered equally influential in identifying the methods selected for a study. The study director was the next most influential party and the board member responsible considered as less so, although their assumed views appear to be factored in earlier. But there are also some signs that the approaches to study-design are changing. The detailed methodology is now the subject of discussions between teams and in-house specialists, and through a process of ‘internal challenge’ of the plans, there is a greater peer review designed to assess whether the methods proposed are likely to allow relevant conclusions to be drawn. Much of the study-design work is bespoke and tailored to the demands of an audit of a particular client and subject matter.

TRENDS IN THE USE OF METHODS 1980s–2000s

A key part of study-design is the choice of methods, to which we now turn. The analysis of the 1980s and 1990s (Table 5.1) is taken from Lonsdale (2000), and is an analysis of information in the methodology annexes of 300 reports published in eight individual years. The more recent data were gathered in 2009 as part of a process improvement review undertaken by staff within the NAO. The data were drawn from a methods database, which also gathered details from the methodology annexes of published

Table 5.1 Methods employed in National Audit Office reports: number of studies making use of each method

Method	1986	1987	1988	1993	1994	1995	1997	1998
Number of studies	30	29	27	39	51	47	38	40
Document examination	30	29	27	39	51	47	38	40
Interviews	30	29	27	39	51	47	38	40
Sample examinations	4	3	13	6	17	4	10	12
Systems/procedures reviews	0	1	0	0	2	1	7	8
Analysis of existing data	9	6	8	5	4	10	12	15
Surveys	0	4	2	14	20	10	12	12
Focus groups	0	0	0	0	1	3	5	4
Expert panels	0	0	0	1	0	4	3	7
Case studies	6	1	6	12	16	6	3	3
Site visits	10	10	7	12	22	13	14	13
Observation	0	1	0	2	0	2	0	3
Deadweight	0	0	0	0	0	1	0	0
Displacement effect	0	0	1	0	0	0	0	0
Correlation	0	0	0	3	6	1	1	0
Sensitivity analysis	1	0	0	0	3	1	1	2
Decision theory	0	0	0	0	0	1	0	0
Modelling	0	0	0	1	1	0	1	1
Comparisons (including international)	2	2	3	6	5	3	3	5
Consultations with others	2	5	8	22	29	22	14	22
Bibliometric analysis	0	0	0	0	0	0	1	0
Literature review	0	0	0	0	0	0	4	2
Benchmarking	0	0	0	0	0	0	6	0
Cognitive mapping	0	0	0	0	0	0	0	1

Source: Lonsdale, 2000.

reports. In all cases the accuracy of the data is dependent on teams including comprehensive details in their published report.

1986–88

In this period we see the overwhelming importance of document examination and interviews for the collection of evidence for VFM studies. Indeed, these appear (although there was often no formal methodology annex to record what was done) to have been the major, if not the only, methods employed in many reports. Discussions with departmental staff,

examinations of samples of transactions, and the review of official files were crucial. In addition, the NAO's other main means of obtaining information were site visits and the examination and (relatively simple) analysis of existing data, perhaps from the department or bodies involved. Questionnaires and international comparisons were infrequently used. On a small number of occasions outside consultants were brought in to help.

1993–95

In this period a number of developments appear to have taken place. There was a continuation of the use of more traditional methods but we also see a noticeable growth in the use of questionnaires and surveys, which became an important feature of NAO work in the 1990s. In 1994, for example, 20 studies published in that year made use of that method. Another feature was the growth of consultation with third parties, with more than half of reports examined making extensive use of consultation with interested parties. Finally, a feature of the second period is what we might call a small-scale, but nevertheless noticeable, increase in experimentation and thus the use on one or two occasions of new methods such as focus groups, decision analysis and modelling, which had not been seen before.

1997–98

In this period, interviews and document examination remained the bedrock of VFM work, but there was greater attention paid to how these methods were employed. Surveys remained another frequently used method – some studies surveyed several groups of people – as were visits to a range of sites to collect evidence. Efforts to consult widely with, and gather evidence from, third parties continued. There was greater use of focus groups and panels of experts compared with the previous period. Benchmarking was seen in a number of studies in 1997, and the period also saw the first use of a number of other methods such as literature reviews and cognitive mapping.

2003–05

During this period there was considerable use of a range of methods that had begun to appear in the 1990s. Focus groups, case studies, forms of international comparison and literature reviews all appeared in around one third of reports, whilst surveys of one kind or another were used in more than half of audits. There was increased interest in the work of

evaluators, reflected in the preparation of a guide on how evaluation methods could be applied to audit (National Audit Office, 2001). The 2000s also saw the growing influence of external links – academic reviewers who commented on draft reports and advised on methodological improvements, expert advisors (used on more than 50 per cent of audits in 2003–04 and 2004–05), and growing groups of suppliers and partners, who provided access to sector and technical knowledge, which was applied to audits (Lonsdale, 2008).

2005–09

Overall, Figure 5.1 shows continued heavy use of interviews and document examination, as well as surveys, stakeholder consultations, quantitative data analysis, site visits, expert panels and focus groups. Literature reviews were undertaken in between one-quarter and one-third of studies, emphasising a willingness to place the study in a wider context which had not been evident ten years earlier. We also see a long tail of methods used less often, including modelling, Geographical Information System (GIS) mapping and process mapping. Our data for this period indicates that auditors used 26 distinct methods, some routinely, others only intermittently. Figure 5.2 suggests that there has been an increase in the average number of methods used per study (from 6 in 2005–06 to 7.5 in 2008–09), with all but one of the six teams examined showing increases, some quite significant (for example, the auditors of Private Finance deals). This may reflect the importance attached to having a strong evidence base and the need for triangulation of sources (National Audit Office, 2008a), which was given a strong push.

Looking back over more than 20 years of value-for-money reports, it is possible to identify a number of trends. In particular:

- performance audit has remained grounded firmly in the basic core repertoire of interviews, document examination, data analysis and surveys;
- there has been an incremental addition of new methods to the core repertoire, especially in the 2000s, so that the performance auditors' 'toolkit' includes around 30 different methods;
- there has been increased use of methods which do not stem from the audit and accounting disciplines, in particular, greater use of qualitative methods such as focus groups and, more recently, for example, decision analytical modelling;
- there has been an increase in consultative methods, where audits have sought to examine the views of service users and other

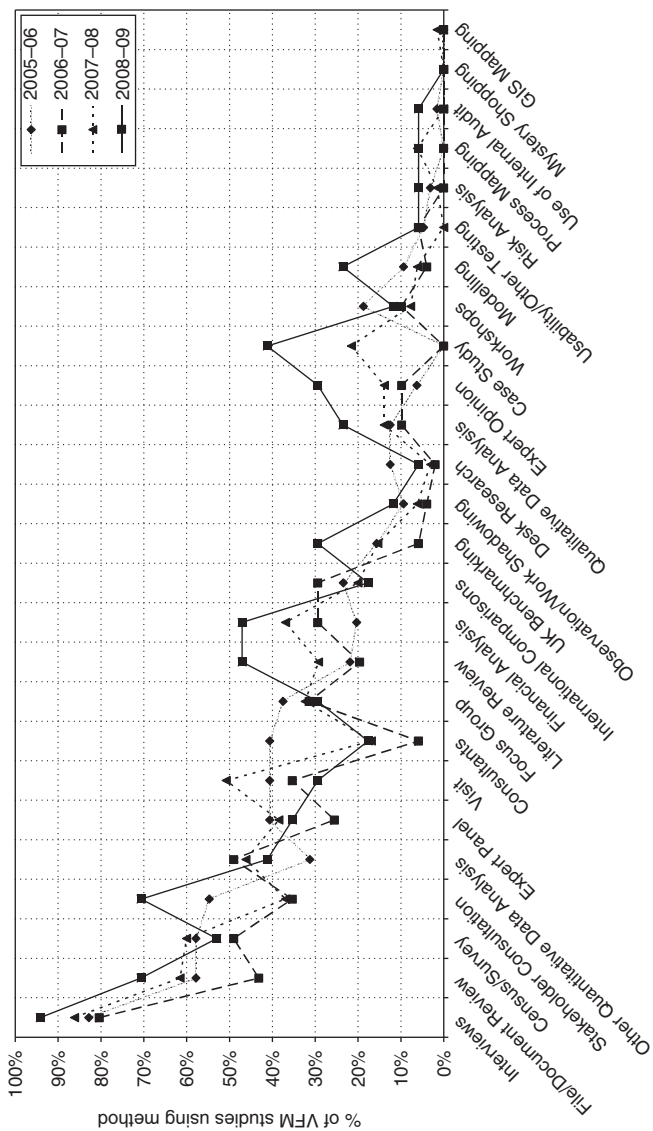


Figure 5.1 Methods used by the NAO 2005-06 to 2008-09

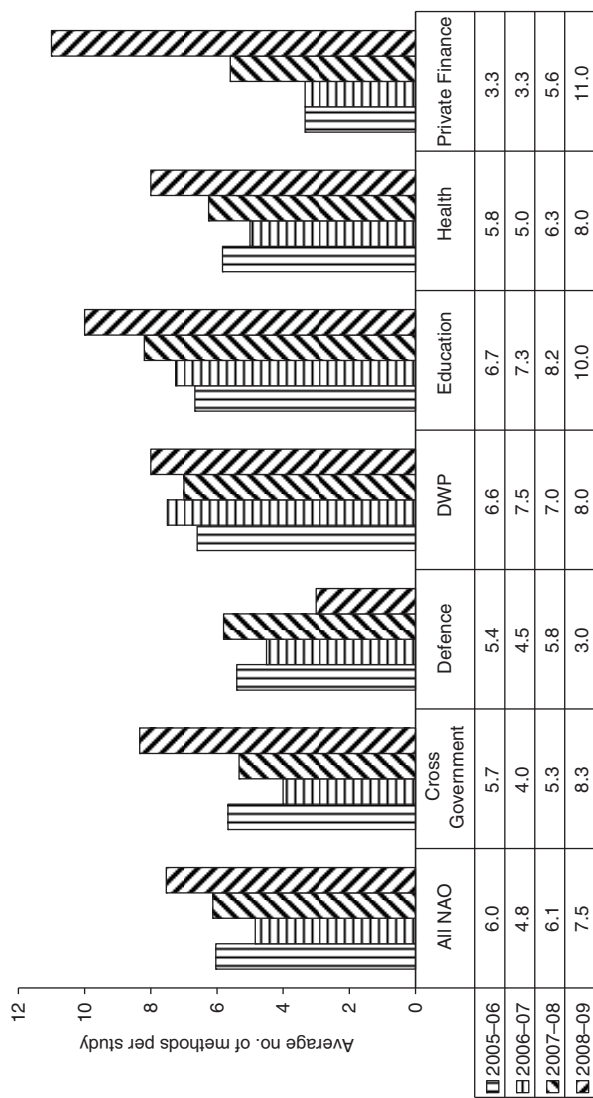


Figure 5.2 Average number of methods used per study 2005-06 to 2008-09

- stakeholders or service providers, reflecting the position within the NAO that it works on behalf of the taxpayer and needs to find ways of building user views into its reports; and
- although we do not have data for the early period, the evidence suggests that many teams are using a slightly larger number of methods on each study than in the past, indicating a recognition of the importance of a convincing evidence base.

FACTORS INFLUENCING CHOICES

We now turn to examine the influences that appear to have shaped the choices summarised above. Research in the late 1990s (Lonsdale, 2000) suggested that the changing environment – political, administrative and intellectual – in which the NAO operated had been the main driver encouraging the use of new methods. This included the growing expectations of audit from those subject to it, and more emphasis within government on achieving value for money. The NAO's interpretation of these changes and the translation of thought into action had come in particular through the lead given by the (then) C&AG, who encouraged innovation. Greater awareness of methodological issues amongst some staff had helped them to respond and there was a growth in guidance setting new expectations for audit work.

An analysis of available literature (Lonsdale, 2000, 2008; Ling, 2007) and the author's short survey of VFM managers responsible for most of the 60 VFM reports published in 2008–09, along with a series of focus groups with VFM staff, helped to identify those factors that appear to influence the choice of methods. These have been brigaded into four groups – those that are specific to the team themselves (team centred); those that are related to the audit project (project centred); organisational influences and finally, wider environmental factors. Each is considered in detail.

Team-Centred Influences

To understand what factors were important at a team level, audit managers within the NAO were surveyed to try to identify what they considered shaped their choices, both in terms of encouraging or restricting them. Managers were asked for their three most important factors, in order (Figures 5.3 and 5.4). In terms of the positive drivers, the most influential factor was the nature of the subject to be examined. Thus, for example, an audit of a health topic may require a survey of hospitals to gather data or a

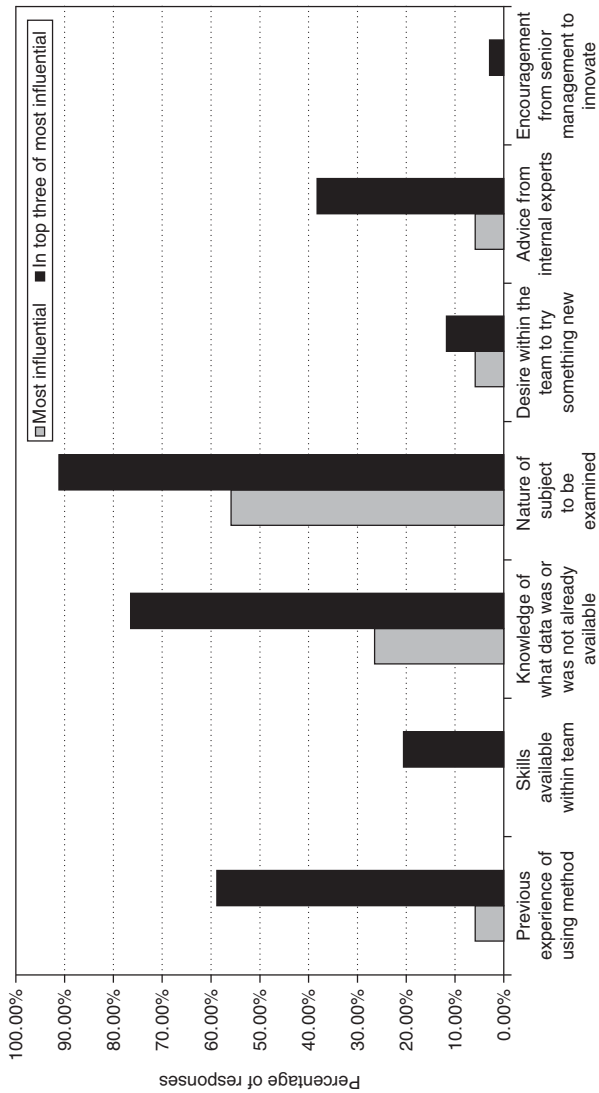


Figure 5.3 Factors managers consider shape their choice of methods

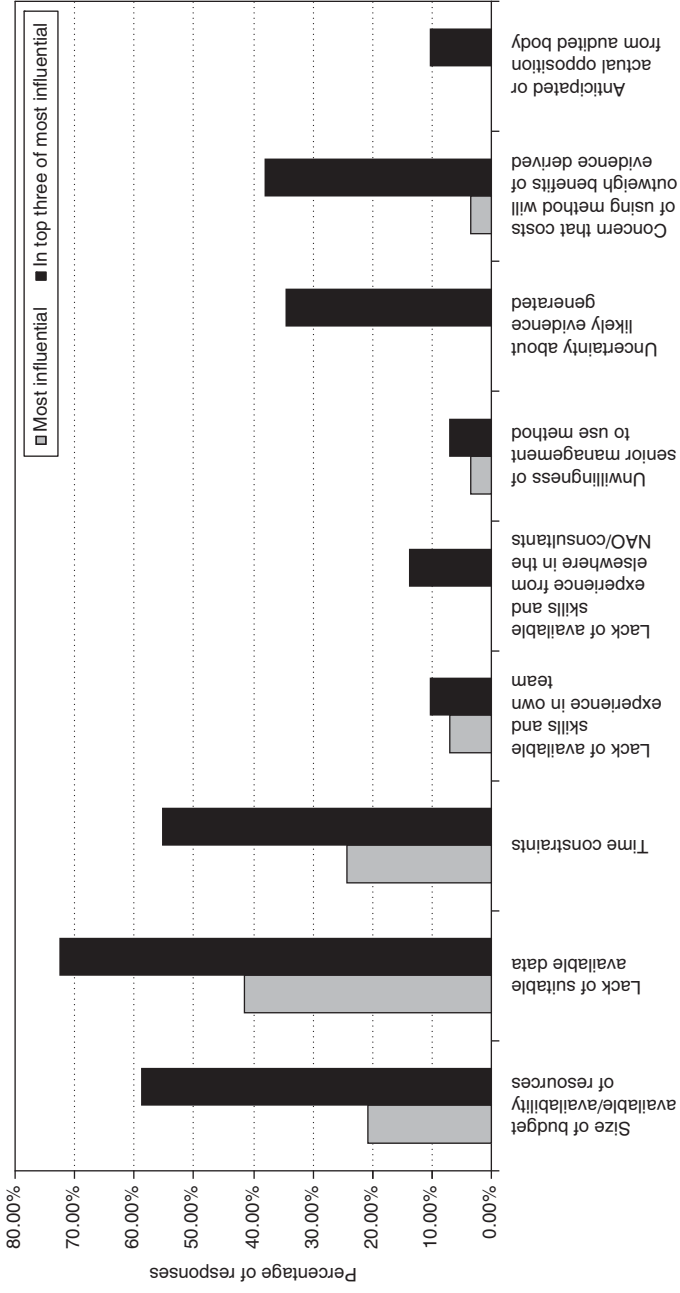


Figure 5.4 Factors managers consider constrain their choices

series of site visits; an audit looking to comment on user views may involve focus groups. The second factor was consideration of what information was, or was not, already available. This included the need to fill gaps in existing knowledge, whereby auditors would rely on the work of others (where they trusted it), or to avoid duplicating what was already known. And the third most important factor was previous experience of having used the method before amongst the members of the audit team. Drawing on past experience can be good in terms of ensuring teams make use of their knowledge and develop an expertise. The downside can be rather predictable work, whereby audits tend to be approached in much the same way.

Managers were also asked about the constraints they felt they faced. Keen (1999) has already noted the tendency for auditors to gravitate towards methods that were acceptable to both auditor and auditee, but by far the most significant limitations were three practical issues – the lack of data to allow them to use particular methods, and the availability of time and of resources. As mentioned earlier, the NAO has sought to produce its reports more quickly, in part to meet parliamentary timetables, and therefore there are limits to the amount of time available. The latest Comptroller and Auditor General (appointed in 2009) has emphasised this even more and in 2010 a number of reports were produced in three months. Thus, there are practical constraints on how extensive or time-consuming the methods used can be within the resources available. Linked to this, two other factors were: uncertainty about the likely evidence to be generated; and concern that costs of using the method would outweigh the benefits. Interestingly though, lack of available skills was not seen as a particular issue; managers apparently feeling able to secure skills when they needed them.

The influence of individuals within teams on methods choice can be seen from the example of the use of decision analytical modelling (Bechberger et al., 2011). Although efforts had been made in the late 1990s to encourage the use of an external panel of operational research experts, little progress had been made in integrating such methods into VFM audit. An important reason for this appears to have been the very limited number of staff within the office able to act as ‘intelligent customers’. Almost a decade later, however, a number of staff were recruited with existing knowledge of, and skills and interest in, decision analytical modelling. In addition, the NAO employed an academic specialist in operational research on an academic secondment. His contribution included developing some basic training courses and guidance to raise awareness, and he also participated in a number of studies.

The health team within the NAO made good use of modelling in a series of reports in 2008 and 2009, helped by a supportive director and by

subjects which lent themselves to using the approach, particularly because data were available. They were faced by a number of problems including scepticism (or fear of scepticism) from audited bodies, the tight deadlines for audit studies and the limitations of a short report in which to explain their methods. But they concluded that these concerns, whilst valid, could be overdone, and that 'Overall . . . Operations Research can be a valuable addition to the auditor's toolbox, providing new insight into problems under investigation. When used appropriately and integrated within a study that utilises a range of other methodologies, models allow the examination of counterfactuals and are a basis for quantifying the outcomes of recommendations' (Bechberger et al., forthcoming).

Project-Centred Factors

A second group of factors relate to the type of audits being undertaken and their subject matter. Many audits are focused on individual organisations. Developments in the public sector over the last 20 years have seen greater diversity in the types of organisations delivering services (Pollitt and Bouckaert, 2004). Auditors have had to respond by looking at different models, and in doing so have had to apply different methods. So, for example, where there is devolved delivery at local level, process mapping has been used to help understand the linkages. The type of questions to be addressed have also shaped methods choice. Where teams are essentially telling a story as to what happened, perhaps when examining a project-management disaster, official documentary evidence and interviews with participants can often provide the bulk of the necessary information. Where the study is considering whether resources could have been spent differently, teams have taken to using modelling. An example of this is the NAO study on 'End of Life Care' (National Audit Office, 2008b), where the modelling work suggested the potential to release substantial resources, whilst delivering care which better met patients' needs by decreasing the utilisation of acute care over the last year of life by cancer and organ failure patients.

The history of the successful use of particular methods in previous studies is another project-centred factor. Teams tend to copy and draw from each other and where they have seen a method used successfully in a particular type of audit have sought to draw on that experience. Finally, there is the nature of the evidence needed for the project. Where projects are considering the views of users or providers of services, surveys feature strongly.

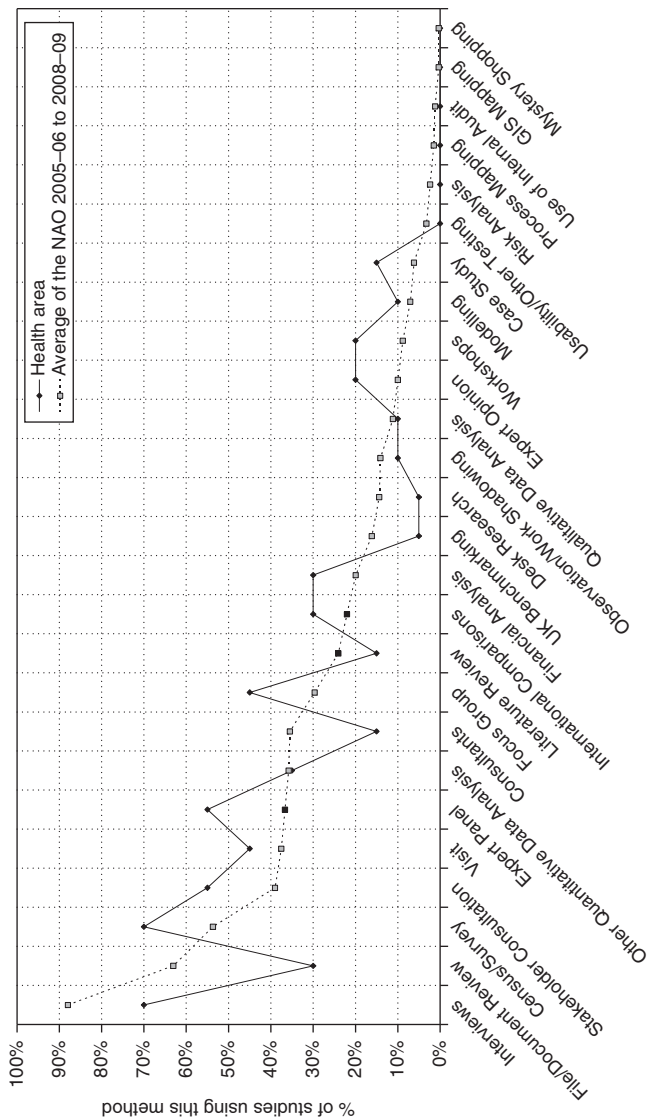
The comparison between two areas of the NAO in terms of choice of methods illustrates that teams undertaking different types of work use

different selections of methods. The health VFM team, for example, have made above-average use of some of the more commonly used methods – surveys, stakeholder consultation, visits and expert panels, for example (Figure 5.5). Many of the methods they use involve consultation in one form or another with stakeholders, reflecting the wide range of interested parties in the NHS and the importance that the teams place on having external credibility amongst professionals. In contrast, the cross-government team (Figure 5.6) – whose work is unusual in not being focused on a single department, but rather looks right across government – have made above average use of methods used less often by the NAO as a whole, including workshops, modelling, process mapping and case studies. The latter, in particular, reflect the types of reports they produce, which have often looked at ‘interesting/good practice’ across government through the use of detailed case studies. They have also brought in specialists to write think pieces, which have been published with reports because there was deemed to be a wide cross-government audience for lessons learned.

Organisational Factors

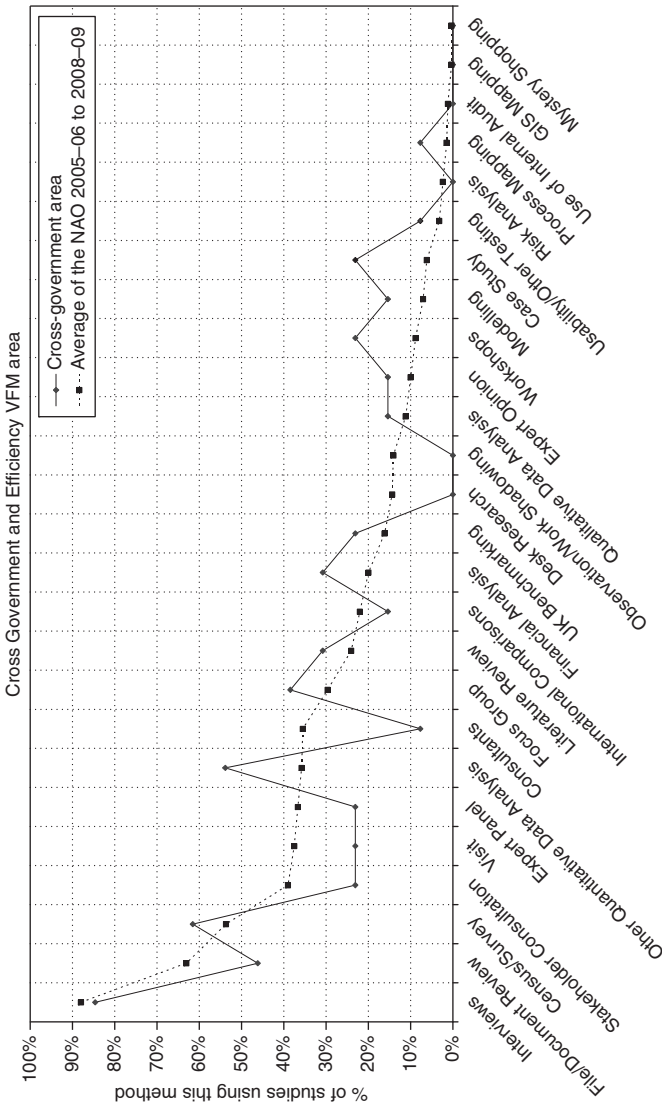
The nature of the NAO as an organisation is also an important driver. VFM audit is undertaken by an audit institution with formal statutory powers and responsibilities to Parliament to do a range of work. Alongside reporting on VFM, for example, the NAO provides an opinion on over 400 sets of financial statements. This creates constraints on VFM (for example, it dictates the amount of money spent on the conduct of VFM work as opposed to other types of audit), sets expectations (for example, what are considered appropriate standards of evidence are rooted in the financial audit tradition, and there are assumptions about the prime significance of formal and official written sources of evidence), and also determines the core skills of many staff (the majority of whom are trained in accountancy). Undertaking VFM audit within a body also carrying out financial audit work encourages an interest in particular aspects of government: the stewardship of resources, corporate governance arrangements, risks, propriety and the regularity of spending, which in turn influence the type of methods used. There have been growing expectations, for example, to look at ways of measuring costs and a strong emphasis on quantification, encouraging an interest in financial analysis.

The nature of the organisation may also encourage what some may see as rather conventional approaches to evidence gathering. Many VFM studies are typically based on interviews with responsible officials and agencies, focus groups, a survey and file examination. As we have seen



Note: Covers studies published up to 25 February 2009.

Figure 5.5 Choice of methods in health studies 2005-06 to 2008-09



Note: Covers studies published up to 25 February 2009.

Figure 5.6 Choice of methods in cross-government studies 2005-06 to 2008-09

there is relatively limited use of more demanding or more time-consuming methods. The move to greater use of qualitative data has had many benefits for the NAO's ability to explore the views of service users and unpack problems, but there appear to be clear limits to the type of evidence that a state audit institution considers appropriate or possible. There has also been no real attempt to undertake longitudinal research, which would require much longer timescales, or try more experimental designs, where the evidence may be less conclusive or seen by those audited as less reliable and convincing.

From an organisational point of view, there may be considerable merit in concentrating on tried and tested methods. Ling (2007) has argued that introducing innovative approaches to performance audit might push existing institutional arrangements to bursting point. He argues that the NAO has responded to pressures on it to enhance its work by including more outsiders to collect and analyse data, has established a more inclusive approach to working with stakeholders and has encouraged more team-working and inter-team collaboration. The introduction and development of external expert assessment into the NAO's processes has also clearly added a new discipline on study teams and strengthened incentives to produce good quality evidence and carefully substantiated and penetrating analyses. Ling argues that these are appropriate responses but suggests that there are risks. The admissibility of new types of evidence may lead to greater challenge to the credibility of audits, and more complex studies raise questions about quality assurance. As he puts it 'The [NAO] is committed to the use of evidence that is relevant, reliable and sufficient. Defining and managing these desirable aims will become harder and more contentious.'

Organisational resource limits within the NAO are always present. The NAO has a limited budget like any public body, but is responsible for auditing central government income and expenditure of over £800 billion a year in a wide variety of fields. To do this it has limited skills available and cannot expect to employ specialists in every discipline. Although it now recruits staff with a wider range of skills into its VFM work, the majority remain predominantly accountancy trained. This has many advantages and provides excellent skills in the core VFM methods, but it does mean that many staff have little or no experience of other approaches, such as cost-benefit analysis or economic and statistical modelling.

Environmental Factors

The NAO operates in a complex and political environment, working under time and resource pressures and influenced in particular by the need to produce sufficient reports for the PAC and to deadlines to meet

the Committee's timetable. It operates in a 'political context', which has a strong influence on methods choice. Its role in examining public policy developments on which senior officials will be questioned gives rise to expectations that the evidence base will include interviews with senior staff, and the review of 'key' documents. There is also an assumption that those to be held accountable for the reported performance will have the chance to comment on the reports. This clearance process (Sharma, 2007) – in which the draft report and possibly the supporting evidence are exposed to those who are the subject of the work to give them the opportunity to comment on the accuracy and completeness of the facts – strongly encourages auditors to have defensible evidence (a point discussed further in the next chapter) and again emphasises the importance of using methods which are convincing for those subject to them. Problems can arise where evidence is weak. Source materials that teams have to work with can be poor. In many departments and agencies, for example, unit cost data is often not available, and time recording is not undertaken. Where data is available, it is often of poor quality or incomplete (such as including direct costs only).

The parliamentary context in which the NAO operates also has an influence on what methods are considered appropriate. Audit reports play an important role in accountability processes and the main audience are parliamentarians. Lonsdale (2000) reported that there was limited interest within the Committee in the way in which VFM was undertaken as opposed to the conclusions that had been drawn. To a great extent, the methods employed were taken for granted. Subsequently, there is some evidence of greater interest in the approaches used, some of it more critical. During one hearing of the Committee of Public Accounts on a report on the value for money of a private health initiative, for example, one Member of Parliament commented (Hansard, 2005):

I am really very uneasy about this Report. . . . When you read the Report carefully it is actually bereft of financial analysis frankly. I read it twice and then I went to the methodology and in the methodology there is only one reference to financial analysis, the rest of it is qualitative analysis based on focus groups, expert panels, case studies and the case studies are just looking at document reviews and interviews with stakeholders. The vast bulk of the work which has been done has been qualitative.

Auditors consider comments from parliamentarians highly important and their concern with increasing the extent to which data analysis featured in reports led to increased training for staff from 2006 onwards. In general, quantitative data appears to be regarded as the most authoritative, in particular that generated independently by the NAO itself. Amongst

the parliamentary committee there can be mistrust of official government data, unless it has been validated by the NAO. Qualitative evidence appears to be regarded as of lesser value, albeit helpful in illuminating particular issues.

Finally, auditors are conscious that although they are independent of government and have statutory powers which allow them to undertake the work as they see fit, they need to secure agreement from those they audit as to the value and merit of the methods to be used. Generally, teams expose their proposed methodology to their counterparts within the body and may take account of suggestions (for example, about site visits). On occasions, there can be some hostility to certain methods. One form of evidence gathered that has been received with some suspicion is mystery shopping.

CONCLUSIONS

This chapter has examined how contextual factors have shaped methods choices for performance audit. We have seen that over time there has been a gradual widening of the methods used by audit teams, through incremental developments though, rather than major step change. Although VFM audit remains strongly rooted in the examination of documents, the use of interviews and surveys, and the analysis of existing data, we also saw the addition of a range of other methods drawn from different disciplines. This emphasises that VFM audit is a hybrid activity which is likely to evolve to meet the needs placed upon it.

Examining the context in which performance audit has taken place is a good way of considering why these changes have occurred. Context is clearly crucial for both driving and constraining choices of methods. There appear to be multiple influences shaping the development of methods within the NAO's performance audit work – at team, project and organisational levels, as well as shaped by the external environment. Individuals will bring a history of previous experience with them so that they have preferences or skills (or lack of skills) which they can bring to bear on a subject. Teams will be affected also by the demands of their particular project and the area they are studying. The fact that they are carrying out the work within an audit office, as opposed to a university department or a research consultancy, will determine expectations around methods, budgets and reporting arrangements, whilst the environment – the accountability setting in which the work is done, for example – may constrain because of the need to produce work convincing to a non-specialist audience. The contextual factors can thus be multi-layered.

Such contextual factors are constantly changing. One consequence for the NAO of the growing concern, initially, for efficiency and more recently about 'value for money' within government has been a renewed emphasis on quantification of conclusions and greater financial analysis. Pollitt et al. 1999 noted a surprising lack of data analysis within many audit reports, as we have seen some parliamentary readers did some years later. Subsequently, the importance of strong data analysis, clearly set out in published reports, has been highlighted by NAO guidance (National Audit Office, 2008a) and this has been pushed even more strongly as the public sector has come under growing pressure for savings.

REFERENCES

- American Evaluation Association (2009), 'Presidential address', available at: www.eval.org/eval2009/ (accessed 14 February 2010).
- Bechberger, E, D Lane, T McBride, A Morton, D Quintas and CH Wong (forthcoming), 'The National Audit Office uses OR to assess the value for money of public services', *Interfaces*.
- Exchequer and Audit Department (1981), 'Review of Value for Money Audit', London: Exchequer and Audit Department, unpublished.
- Flint, David (1988), *Philosophy and Principles of Auditing – An Introduction*. London: Macmillan.
- Hansard (2005), 'Committee of Public Accounts, Minutes of Evidence', hearing 17 October 2005, Q62, available at: www.publications.parliament.uk/pa/cm200506/cmselect/compubacc/562/5101705.htm (accessed 25 May 2009).
- Keen, Justin (1999), 'On the nature of audit judgements: The case of value for money studies', *Public Administration*, 77 (3), 509–25.
- Ling, T (2007), 'New wine in old bottles? When audit, accountability and evaluation meet', in Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (eds.), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Jersey: Transaction Publishers.
- Lonsdale, Jeremy (2000), 'Advancing beyond regularity: developments in VFM methods at the National Audit Office 1984–1999', Unpublished PhD thesis, Brunel University.
- Lonsdale, J (2008), 'Balancing independence and responsiveness', *Evaluation*, 14 (2), 227–48.
- Lonsdale, Jeremy and John Mayne (2005), 'Neat and tidy . . . and 100% correct: assuring the quality of supreme audit institution performance audit work', in John Mayne and Rob Schwartz (eds.), *Quality Matters: Seeking Confidence in Evaluating, Auditing and Performance Reporting*, New Jersey: Transaction Publishers.
- National Audit Office (2001), *Evaluation and its Use in Value for Money Audit*, London: National Audit Office.
- National Audit Office (2008a), *VFM Handbook*, London: National Audit Office.
- National Audit Office (2008b), *End of Life Care*, London: National Audit Office.

- Pollitt, Christopher and Gert Bouckaert (2004), *Public Management Reform: A Comparative Analysis*, 2nd edition, Oxford: Oxford University Press.
- Pollitt, Christopher and Hilkka Summa (1996), 'Performance auditing: Travellers' tales', in Eleanor Chelimsky and William Shadish (eds.), *Evaluation of the 21st Century: A Resource Book*, Newbury Park, CA: Sage.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilkka Summa and Marit Waerness (1999), *Performance Audit and Public Management in Five Countries*, Oxford and New York: Oxford University Press.
- Sharma, N (2007), 'Interaction and interrogations: Negotiating and performing value for money reports', *Financial Accountability and Management*, **23** (3), 289–311.

6. Evidence and argument

Alex Scharaschkin

It is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits; it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician scientific proofs. – Aristotle, *Nicomachean Ethics*, i.3. (trans. W.D. Ross)

INTRODUCTION

Chapter 5 examined the selection of methods to generate evidence for performance audits. It showed that evidence considered adequate by auditors can be derived in a variety of ways and that there has been an expansion over time in the methods used by auditors. This chapter examines how auditors use the evidence they gather to reach conclusions. Such conclusions are put forward on the basis of a process of argument founded upon the evidence.

Much of the ‘art and craft’ of performance auditing is concerned with building up this argument, which must be rigorous, objective and fair (see Lonsdale, 1999; Keen, 1999, Furubo in this volume). Yet, although auditors may use scientific studies and academic research, performance audits are not themselves scientific studies or academic research projects. Moreover, they must be delivered with limited resources to tight time-scales and for this reason, auditors find themselves periodically reflecting on what constitutes a sufficient and complete argument in audit. Making the kinds of argumentation used in performance audit work more overt, and considering them formally, also allows auditors to reflect on, and hence improve, practice.

In this chapter I consider the nature of evidence and argument in performance audit, drawing on experience at the National Audit Office and its particular variant – value for money (VFM) audit. I suggest that while the *methods* of analysis and interrogation of evidence used in this kind of work are for the most part those used in social science research, the structure of argumentation is more akin to that required to support a legal judgement.

Adopting this view explicitly in designing studies and critiquing emerging findings can help auditors to assess the sufficiency and completeness of their evidence base and the strength of their conclusions. It can also help those using such reports to understand how they should be employed in supporting accountability and more generally improving public services.

In carrying out their audits, study teams are usually concerned with making inferences from a mass of (possibly contradictory and often incommensurable) evidence, in the context of a sometimes adversarial relationship with audited bodies. This is in contrast to evaluators, who are generally, although not always, contracted at the request of the clients they are evaluating. Work on the nature of evidence and argument in legal reasoning can be usefully applied to gain a better understanding of how conclusions of performance audits are supported by the material included in reports.

Particularly relevant is recent work on using so-called ‘non-monotonic’ logics to model the ‘defeasible’ (defined as capable of being annulled or invalidated) argument patterns that are characteristic of legal reasoning on the basis of evidence (for example, Prakken, 2004a; Bex et al., 2003). Drawing on such work may help to elucidate what makes for a sound argument in a performance audit study, by providing both a ‘language’ for describing current practice, and possibly standards for good practice. It may also provide a counter to the demand of some stakeholders that the audit report should deliver a simple answer to the complex question – ‘Did it work?’. This is because, as set out in the next section, findings in studies are not usually adduced on the basis of purely deductive logic. Reflecting on these issues helps to shed light on current audit practice, and suggests areas where there may be potential for further development and enhancements in audit methodology. This chapter concludes by suggesting activities that could help develop approaches to the analysis and use of performance audit evidence.

A FRAMEWORK FOR ASSESSING AUDIT ARGUMENTS

The quotation from Aristotle’s *Ethics* that serves as a motto for this chapter summarises its key concern. How can auditors ensure that their arguments are constructed with a degree of precision that is appropriate to their subject-matter, their context, and their intended purpose (especially when these aims can pull in different directions)?

This is a question about the relationship between the evidence that has been examined in the audit, and the conclusions that have been drawn.

When, in this chapter, I use the term ‘conclusions’, I mean the ‘findings’ or assertions in audit reports: the descriptions of what is deemed to be the case, based on the audit evidence. In value-for-money reports, these are essentially the auditor’s assessments of ‘the value’ (that is, the outcomes) achieved by the programme, organisation or intervention being audited, and of ‘the money’ spent to deliver that value. Together, these are necessary, but not sufficient, to arrive at a conclusion on value-for-money – although they may be sufficient to arrive at a conclusion on performance in a traditional performance audit.

Concluding on value-for-money, however, involves not only establishing what was the case, but also evaluating whether the value that was delivered was ‘good’ (or ‘poor’, or perhaps even ‘optimal’), given the money that was spent. This requires assessing the extent to which outcomes *could reasonably have been different*, by locating the actual configuration of value and money that was delivered within a wider ‘space’ of ‘reasonably possible’ configurations. These reasonably possible configurations are the (counterfactual) levels of value achieved and money (spent) that could have arisen had other approaches to delivering the programme or service been used and other outcomes achieved.

There are a number of tools that auditors can use to do this. For example, the wider outcome space can be modelled. Outcomes (either actual, or counterfactual) are paths in this space (that is, combinations of money and characteristics of value that change over time), and it is then possible to compare actual outcomes with ‘reasonably possible’ alternatives. For further discussion of comparing actual and counterfactual outcomes in VFM audit, a consideration of which would take us beyond the scope of this chapter (see Bechberger et al., 2010; Scharaschkin, 2010).

Leaving aside, then, the issue of how to form an overall judgement on value for money from the individual conclusions of the audit, I argue that the test of whether a study is sufficiently ‘precise’, in Aristotle’s sense, is whether its conclusions are supported rigorously with respect to a suitable *defeasible argumentation system*. By defeasible argumentation I mean reasoning, the validity of which depends on the evidence available: the discovery of new evidence may *defeat* an argument that was previously sound (or, indeed, may *re-instate* an earlier argument that had subsequently been rejected). The classical example is inductive reasoning (inferring a general conclusion on the basis of particular observations).

In the case of performance audit, complete and perfect data is rarely available, and claims are not advanced with certainty, but rather as rational and reasonable on the basis of what is known to be the case. Performance auditors aim to draw conclusions that present a balanced and fair view given the evidence available; but these may be subject to

revision in the light of new information or changes in circumstances. Defeasible logic provides a natural framework in which to locate this kind of discourse.

However, the criteria for the appropriateness of audit conclusions extend further than purely logical considerations. This is implied in the traditional definition associated with financial audit: to provide a 'true and fair view' of the matter. One could argue that an audit finding is *true* if it is the conclusion of an argument that is valid with respect to a particular logical framework. But this would not necessarily entail its being *fair*. For example, a National Audit Office report on *Managing Attendance in the Department for Work and Pensions* (National Audit Office, 2004) found that many basic procedures in the Department's policy on staff attendance management – such as issuing warnings – are implemented inconsistently or not at all at local level. Although this conclusion is true, it would not be fair to state it as the overall conclusion of the argument in the report, because this study also found that the Department has strengths on which to build, including a well-designed policy, high levels of staff commitment and the experience of many good managers at a local level. If we let P be the statement 'many basic procedures are poorly implemented', and Q be the statement 'the Department has strengths on which to build', then both of the arguments:

' P ; because P and Q '; and ' Q ; because P and Q '

are logically valid. However, unless there is a good reason for weighting the importance of one of the findings P or Q much more highly than the other, neither of these two arguments is fair as the final stage in the chain of reasoning in the report, as they do not present a balanced view of the Department's actions.

So I suggest that a useful conceptual framework for reflecting on the nature of evidential support for performance audit conclusions is a system consisting of:

- (1) a set of *logical* rules including, as a subset, the usual rules for deductive inference (such as ' P ; P implies Q ; therefore Q '), but extending them to cope with the fact that (outside pure mathematics) conclusions with any significant information content are almost never arrived at purely deductively; and
- (2) additional *domain-specific* rules, which, in the case of performance audit, may include particular 'argument schemes' that recur regularly in analysing audit evidence (such as ' E claims X ; E is an expert; therefore [prima facie] X '), as well as a priori principles such as the

requirement that conclusions in audit reports should be balanced in their presentation of the facts.

The domain-specific rules, for example around balance and fairness, reflect the context within which performance audit is conducted and which shapes what comes to be regarded as acceptable reporting. These are likely to reveal different national cultures and institutional structures.

In this chapter I refer to such a conceptual framework as a *defeasible argumentation system*. I do not include considerations such as feasibility or acceptability within the domain-specific rules: these are considered separately later. I develop the concepts introduced in requirements (1) and (2), and illustrate their applicability to VFM audit with examples from the National Audit Office. Key issues that emerge are that:

- (a) it is important to consider explicitly the ways in which lines of argument can be both 'defeated' by evidence, or 'undercut' by attacks on the warrants (the underlying justifications) for inferences made, and this has implications for the way in which challenge and review arrangements are brought to bear on emerging audit findings;
- (b) it is also important to accommodate the fact that, in general, both the arguments that support a particular conclusion and the arguments against it will be of different strengths, depending on the reliability of the evidence and on the warrantability of the inferences drawn from that evidence. Audit reports must reconcile these to present a balanced description of the audit findings;
- (c) warrants (or justifications) for inferences derive both from particular patterns of argument that arise in drawing conclusions on the basis of audit evidence, and from general principles or standards that apply to audit work in general. There is scope for developing a more detailed framework for warrantability in VFM audit, to provide a tool to help assess rigour when developing conclusions following fieldwork, and some suggestions for doing so are given below.

A TYPOLOGY OF KNOWLEDGE FOR VFM AUDIT

As well as defending the proposition that different frameworks for rigour are appropriate in different domains of enquiry, Aristotle in his *Ethics* distinguishes three types of knowledge that enquirers can bring to bear on an issue. *Episteme* is pure, universal knowledge, as of logic or mathematics. *Techne* refers to the particular skills and knowledge required for a specific craft or activity. And *phronesis* (which is often translated as 'prudence',

but has not given rise to any contemporary English term in the way that ‘epistemology’ and ‘technology’ have derived from *episteme* and *techne* is concerned with how to take practical action by being able to ‘deliberate rightly about what is good and advantageous’ (*Nichomachean Ethics*, 1140a).

This threefold typology provides a convenient way of structuring a discussion of evidence and argument in performance audit. I begin with some remarks on the nature of sound argument when reasoning about evidence, and discuss some of the general principles of defeasible logic that underlie such arguments: the *episteme* of performance audit. Then under the heading of *techne* I consider some of the specific attributes of performance audit work that distinguish it from other domains that have similar logical bases, such as evaluation and jurisprudence, and discuss the implications of these for building appropriate levels of soundness and rigour into the audit process. Finally, I acknowledge in considering *phronesis* that the aim of performance-audit work is to bring about beneficial change in the stewardship of public funds and the delivery of services by the public sector, not simply to provide a well-argued case for a set of findings. Although sound argument and compliance with appropriate technical principles when carrying out performance audit is a necessary condition for furthering this aim, it is not a sufficient one.

THE EPISTEME OF RIGOUR IN PERFORMANCE AUDIT ARGUMENT

Deductive and Defeasible Logics

For many people the paradigm of sound reasoning is deductive logic. A deductive argument consists of a collection of premisses and a conclusion drawn from them in accordance with a set of inference rules. A deductive argument is sound if the premisses with which it starts are all true, and each argument step is a valid inference with respect to the admissible inference rules (Box 6.1).

Performance auditors draw conclusions, and subsequently may make recommendations, by arguing on the basis of audit evidence. The validity of these conclusions, therefore, depends on both the soundness of the arguments constructed on the basis of the evidence, and the reliability and completeness of the evidence itself.

The conclusions of a process of argumentation can be warranted or underpinned by ‘nature’, or by ‘substance’.¹ Assertions are warranted by nature if the chain of reasoning that supports them is ‘logically correct’.

BOX 6.1 DEDUCTIVE ARGUMENTS

Example 1: Sound argument

Premiss 1. A deal is value for money if, and only if, certain defined outcomes are achieved before, during and after the contract is let.

Premiss 2. For this deal, the defined outcomes were met before, but not during or after, the contract was let.

Conclusion. This deal was not value for money.

Example 2: An argument which is logically valid, but unsound because not all the premisses are true

Premiss 1. No statistic published by the government is reliable.

Premiss 2. This piece of audit evidence is a statistic published by the government.

Conclusion. This piece of evidence is unreliable.

Example 3: Valid argument (by the logical rule of *modus tollens*²); soundness depends on truth of premisses

Premiss 1. If a complaint is received by the department, it is automatically put on the file.

Premiss 2. There is no complaint on the file.

Conclusion. No complaint was received by the department.

For instance, the conclusion of Example 1 in Box 6.1 is warranted by nature. Assertions are warranted by substance if their support derives in part from specific standards or requirements of the field in which the claim is being made. For instance, the legal standard of proof in civil cases is that the conclusion is true ‘on the balance of probabilities’. A requirement in performance audit work is that conclusions should be balanced in their presentation of the facts.

Unfortunately, deductive reasoning on its own is almost never sufficient to establish conclusions on the basis of audit evidence. This is because deductive reasoning is *monotonic*. That is to say, adding more premisses and inference steps to a valid argument may lead to additional valid conclusions, but cannot affect the validity of existing conclusions: the set of true conclusions can only grow as more information is added.

Most actual reasoning outside mathematics, however, is *non-monotonic*, in that the truth of propositions may change as new information becomes available. The classical example is *inductive* reasoning, in which general conclusions are drawn from (necessarily) limited information. For instance

it was believed in England until 1788 that ‘swans are white’ was a true proposition, on the basis of induction from observed data. But this conclusion was shown to be false with the discovery of new evidence (black swans in Australia). Inductive reasoning is an example of *defeasible* reasoning – the conclusions of an inductive argument may be *defeated* by the discovery of new information.

There has been considerable research activity in the field of non-monotonic and defeasible logic over the last twenty years (see Prakken and Vreeswijk, 2002, for a survey), and also in its application to legal argument (for example, Bex et al., 2003; Prakken, 2004a, b; Verheij, 2003). The epistemological groundwork for much of this work is due to the American philosopher John Pollock (1987, 1995, 2002). Pollock’s system includes the usual rules for deductive inference, but also incorporates a number of defeasible inference rules, called ‘prima facie reasons’. In Pollock’s system, prima facie reasons are general epistemic principles for obtaining beliefs (or drawing conclusions) from other beliefs and perceptual inputs, such as memory, statistical reasoning and induction.

Challenging Audit Conclusions: Argument Games

The only way to attack a valid deductive argument is to attack one or more of its premisses. For example, the second argument in Box 6.1 can be refuted by adducing at least one reliable statistic published by the government. However a defeasible argument can be attacked in two ways. One way is to rebut it with an argument for the opposite conclusion. The other way is to undercut it with an argument for why a prima facie reason does not apply in the given circumstances. This form of attack does not argue that the attacked conclusion is false, but only that it is not sufficiently supported by its premisses. In other words it attacks the connexion between the premisses and the conclusion, rather than the conclusion itself.

If, as I suggest, defeasible reasoning provides a suitable framework for argument construction in performance audit, then clearly it is important to pay just as much attention to the way in which conclusions can be defeated or undercut as to how they can be supported. Prakken’s (2004b) notion of an *argument game* is a helpful way of thinking about this and, as explained below, is to some extent already represented in the ‘clearance’ mechanism in the NAO’s VFM audit practice. Prakken defines an argument game as follows (2004b, 9):

An intuitive way to define the defeasible validity of arguments is in the form of an *argument game* between a proponent and an opponent of an argument. Proponent starts the game with the argument to be tested and then the players

take turns, each attacking the preceding argument. Opponent's arguments must be at least as strong as their targets while proponent's arguments must be stronger than their targets. A player has won if the other player has run out of moves. Now an argument A is defeasibly valid if the proponent has a winning strategy in a game beginning with A , i.e. if he can make the opponent run out of moves no matter how she plays.

This notion captures the idea that, in constructing arguments (drawing conclusions) on the basis of a mass of evidence, it is often possible to find both chains of reasoning that support a given proposition to some extent, and chains of reasoning that refute it to some extent (see the example in Box 6.2). A consequence for auditors is that, once an argument for a conclusion that is critical of an aspect of an audited body's business is formulated, the auditee will naturally respond with an argument for why the conclusion should not be as critical as it is. The process of arriving at final, agreed, and robust findings in a VFM audit report, therefore, requires determining, given a set of conflicting arguments, which (if any) of those arguments prevail.

Modelling argument construction in performance audit as a process of defeasible reasoning is useful because it can help to suggest standards for comparing conflicting arguments to see which, if any, is stronger than the other. These standards arise partly from the (domain-free) inference rules of the defeasible logic underlying the process. For instance, each of Pollock's *prima facie* reasons comes with specific associated undercutters, which describe what would be required to undercut a conclusion drawn on the basis of that reason. But there are also domain-specific modes of reasoning that provide (defeasibly) valid warrants for conclusions in performance audit. Such so-called *argument schemes* form part of the *techne* of VFM audit, to which I now turn.

THE TECHNE OF VFM ARGUMENT CONSTRUCTION

Argumentation Schemes

Consider the following statements. (Statements (1) and (2) are taken (slightly modified) from Verheij, 2003.)

- (1) Person E says that P . Person E is an expert with respect to the fact that P . Therefore P .
- (2) Doing act A contributes to goal G . Department D has goal G . Therefore Department D should do A .
- (3) When interviewed, official O in Department D said that Q . Therefore Q .

BOX 6.2 AUDIT EVIDENCE MAY TEND BOTH TO SUPPORT AND TO REFUTE A PROPOSITION

The NAO's report on *Reducing Vehicle Crime* (National Audit Office, 2005) concluded that 'closed circuit television cameras [in car parks] have reduced vehicle crime, although performance varies'. One source of evidence for this conclusion was interviews with car-park operators. Six of the operators interviewed considered the installation of cameras to have been their most successful initiative in reducing crime. However, three operators in the same series of interviews thought that it had been the least successful.

This evidence suggests that there are *prima facie* reasons to support both the statements 'CCTVs have been successful as a way of reducing crime', and 'CCTVs have not been successful as a way of reducing crime'. To determine which conclusion (if either) should prevail, the study team had to assess other evidence. They commissioned a literature review of existing evaluations of crime reduction initiatives, which suggested that the use of cameras can lead to a reduction in crime in car parks (though the extent of reduction varies); and also examined a Home Office evaluation of the impact of CCTV, which showed reductions in vehicle crimes following the introduction of CCTV cameras in a number of case studies. This additional evidence allowed the team to arrive at the conclusion quoted.

All three statements express patterns of argumentation. The first is a version of *argument from expert opinion*, and the second is an example of *means-end reasoning*. Statement (3) bears some similarity to statement (1), but is a pattern one might find specifically in an audit report – we might call it *argument from auditee claim*.

These statements have a number of features in common. Firstly, they all express reasons for conclusions which, it is quite possible to imagine, might be perfectly convincing for particular cases of *E*, *A*, *D*, *G*, *O*, *P* and *Q*. Secondly, it is also easy to imagine scenarios in which these reasons would not be compelling. When assessing these arguments, we ask: under what circumstances would they fail to convince us? In other words, each of statements (1) to (3) has an associated list of critical questions, which explore the extent to which the 'rule of thumb' expressed in the general schema is applicable in an instance of interest.

For example, some of the critical questions attaching to statement (3) could include:

- (a) Is *O* in a position to know *Q*?
- (b) Does *O* have any conflicts of interest with respect to *Q*?
- (c) Is *O* technically competent to understand *Q*?
- (d) Is there any other evidence for or against *Q*?

A third feature shared by statements (1) to (3) is that, formally, they look like *prima facie* reasons of the kind described earlier, although they provide patterns for drawing (defeasible) inferences in more limited domains of discourse (in this case, domains that include ‘experts’, ‘departments’ and ‘officials’) rather than encoding more general modes of reasoning.

All three statements are examples of *argumentation schemes* – patterns of reasoning which are similar in form to defeasible inference rules, in that they can be cast in the form of a set of premisses leading to a conclusion, but are restricted to particular domains of discourse and may only be valid within those domains. The critical questions attaching to an argumentation scheme may function like undercutters of a defeasible rule, or may serve to question the validity of premisses or conclusions, or the conditions of use of the scheme. Argumentation schemes are formalisations of the rules that ‘glue together’ the lower-level conclusions that result from audit evidence, in order to arrive at higher-level audit conclusions (and hence, recommendations).

Figure 6.1 shows an excerpt of the structure of the NAO report on *Reducing Vehicle Crime* referred to in Box 6.2. The arrow marked with an asterisk is the derivation of a conclusion from premisses that is discussed in more detail in Box 6.2. What argumentation scheme is being used here? We can perhaps distinguish two. Firstly, a Home Office study and an independent academic literature review suggest there is evidence that CCTV reduces vehicle crime. This appears to be an *argument from expert opinion*.³ Walton and Reed (2003) examine this argumentation scheme in some detail, and suggest that to evaluate its use it is necessary to address critical questions about the credibility of the expert(s); whether they have appropriate expertise for the field in question; what precisely they asserted that implied the conclusion; whether they are personally reliable; and whether the conclusion is consistent with what other experts assert. These are the sorts of questions the NAO study team will have considered (implicitly or explicitly) in considering the conclusion as admissible on the basis of the research evidence.

Secondly, there is an argument from an *interviewee claim*, possibly combined with a (weak) argument by *induction* (as noted in Box 6.2, more

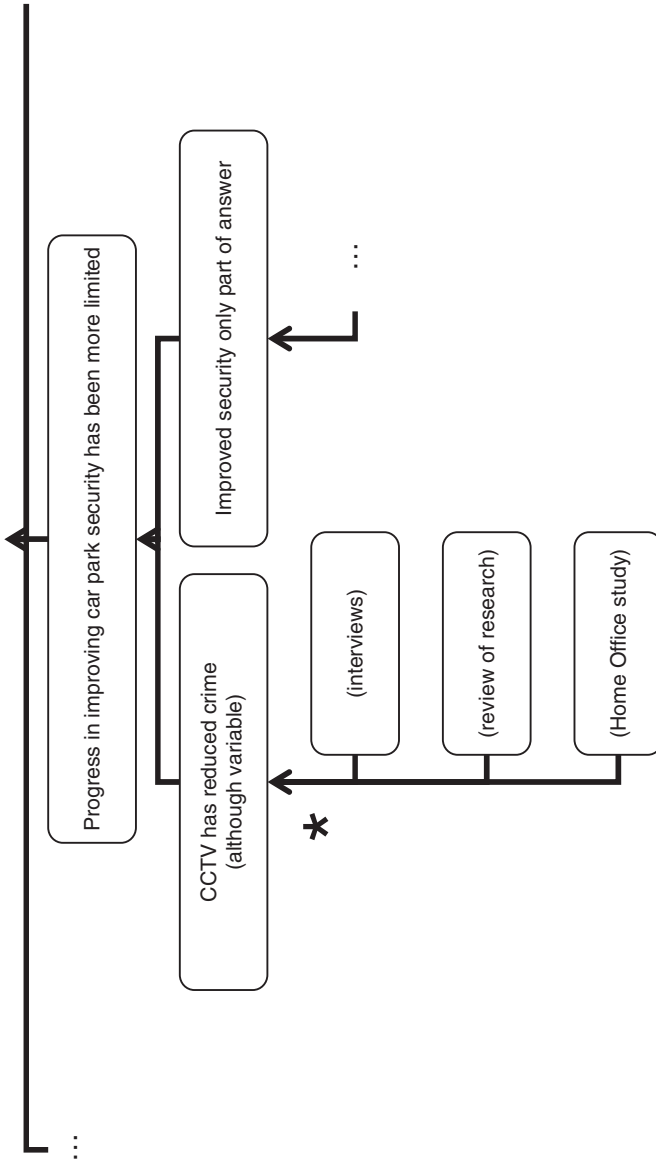


Figure 6.1 Part of a report structure, showing where inferences can be tested

interviews provided evidence for the conclusion than against it). The use of audit interviews as an evidence source has not specifically been considered in the literature on argumentation schemes, and there is no pre-existing set of critical questions here, although the study team will naturally have considered the extent to which interviewees are impartial and in a position to know, as well as more technical (research methodological) issues such as the design of the interview scheme, wording of questions, and the representativeness of the sample interviewed.

As this example illustrates, identifying an appropriate argumentation scheme and set of critical questions is not always straightforward, as to date there has been no research on the types of schemes found empirically in performance audit, and the potential threats to defeasible validity entailed by each. In practice, quality controls such as management and peer review, the reactions of expert panels, and clearance of reports with audited bodies are the usual means by which the rigour of report findings is tested. Further thinking on argumentation schemes in VFM audit could lead towards defining a list of admissible schemes and their associated critical questions.

Using such a framework in practice entails an iterative approach to evidence structuring. As assertions are made by auditors on the basis of evidence, they have to be subjected to the appropriate critical questions. This may lead to the assertions being modified or withdrawn, or indeed to previously withdrawn assertions being reinstated. It may also lead to considerable re-ordering of findings by auditors and their positions within the hierarchy of the audit's argument structure. It may require the collection of more evidence. All are aspects of how auditors refine their thinking as they draw their work to a close.

The use of an *assertions matrix*, a spreadsheet in which emerging assertions are entered, cross-referenced to the different strands of the audit methodology used to address the questions identified in the issue analysis, and hyperlinked to documentation of the audit evidence, can facilitate this process. The rows of the matrix are headed by assertions (emerging findings, given the evidence collected to date). The column headings are the methodology strands (for example, survey results; interview findings; economic analyses; literature review, etc). Entries in the cell corresponding to assertion *A* and methodology strand *M* summarise what evidence from *M* pertains to assertion *A*. Clearly there can be no empty rows (for an empty row would entail an assertion for which there is no evidence), nor should there be any empty columns (for that would mean fieldwork has been carried out that is not relevant to any finding). Potential objections to assertions can be added explicitly to the matrix, in a column adjacent to the assertions, to assist in determining which conclusions eventually prevail as

rigorous and fair representations of the facts of the matter. Experience at the NAO has shown that the assertions matrix approach can help check the evidential rigour of conclusions and, by requiring a continuous revision of the assertions, including consideration of their dependencies, assist in constructing the argument that underpins the audit.

Explicit consideration of critical questions in this iterative way also helps to mitigate the risk of 'confirmation bias' when clearing – discussing the findings and their presentation – with the audited body. As noted above, the NAO's clearance process functions in some ways like a formal 'argument game' to test the strength of the arguments put forward in the report. However, it is natural that the audited body should devote more energy to critiquing those conclusions which are critical of its practice, than those which are not. Having a framework for also subjecting these favourable conclusions to rigorous challenge can help to demonstrate explicitly that they are as robustly reviewed as those that are less favourable to the audited body.

Weighting Evidence and Audit Methods

In the interests of brevity the preceding discussion has passed over a very important point: how the strength of arguments should be determined. It is in the nature of defeasible reasoning that premisses do not support conclusions absolutely, but only to a certain degree. In considering whether one argument prevails over another, how can we compare the degrees of support each lends to its conclusion?

Quantifying evidential support is a topic highly developed in its own right (for an excellent survey see Schum, 1994). Perhaps the best known approach is the Bayesian theory, according to which it is possible to attach to any proposition an a priori measure of one's degree of belief in it. This is then updated, as evidence becomes available (using a mechanism known as Bayes' Theorem), to provide a so-called posterior degree-of-belief measure that takes account of both prior belief, and of the actual data collected. Another approach, which includes the Bayesian theory as a special case, is the Dempster-Shafer theory of evidence (see Shafer, 1976). This provides a formalism for combining evidence from different sources (some of which may be contradictory, and about some of which we may need to withhold belief). Both of these approaches, however, require an initial assignment of numerical degrees of belief to propositions, which is often (but not always) difficult in the context of performance audit. It is beyond the scope of this chapter to examine the merits of these or other approaches to assessing the strength of evidential support, but this is another important area of research where there is scope for considerable development. For example,

research here could help assess the added strength conferred to arguments by ‘triangulating’ audit evidence (that is, using several independent evidence sources to address an audit issue). It could also help in considering the extent to which higher-level conclusions are balanced representations of lower-level findings (for example, a summary conclusion derived from a strongly-supported positive finding and a weakly-supported negative finding should presumably not give equal weight to each).

Finally, although this section has discussed the *techné* of argument construction in performance audit, it has hardly addressed at all the techniques or technical methodologies performance auditors use to collect and analyse evidence: namely, many of the standard qualitative and quantitative techniques of social science research. Some of these were discussed by Lonsdale in this volume. But being able to address the critical questions that attach to the types of argumentation schemes that arise in audit work often depends on having appropriate expertise in these areas. For example, survey evidence used in an inductive argument may be attacked as being unrepresentative, or subject to too high a degree of sampling error, or wrongly extrapolated. Performance auditors must have, or be able to access, the skills needed to address such issues. Rigorous performance audit combines social science research methodology with an approach to drawing conclusions on the basis of findings similar to that required for making a legal judgement.

PHRONESIS

Prompted by Aristotle’s observation that different conceptions of rigour are appropriate in different domains, this chapter has suggested that a concept that can help assess rigour in performance audit is the notion of a defeasible argumentation system. In such a system, the logical connexions between evidence and conclusions arise through valid (though necessarily defeasible) arguments and appropriately used accepted argumentation schemes that have been reviewed in the light of relevant critical questions. In other words, an audit’s conclusions are rigorous if they have emerged victorious from a formal ‘argument game’ of the sort described above.

There are clearly similarities between describing a defeasible argumentation system for performance audit and setting out standards for performance audit practice. Standards (such as the International Standards on Auditing for financial audit, or the US Government Accountability Office’s *Government Auditing Standards* that cover both financial and performance audit) have both normative and empirically derived content. That is, they set out both what *should* be the case (for example to ensure accountability), and what *is* best practice. Once ‘best practice’ is

determined, however, it too defines norms. In other words, auditing standards are informed both by societal expectations, as well as what is actually useful, feasible and successful in practice.

CONCLUSION

This chapter has shown how defeasible logic can serve as a tool for structuring the empirical description of audit practice and hence, potentially, lead to a set of audit standards that encode both external normative content and an appropriate conception of rigour in audit. The question of what external norms could or should be imposed depends on the wider context of audit and accountability, as does the question of where audit effort should be focused, and the role of the audit office in leveraging change through its work. ‘Prudent choice’ – *phronesis* – is required in response to these questions; but this takes us beyond the realm of the technical logic and argument of audit.

NOTES

1. This classification of warrantability draws on Toulmin’s (1958) well-known analysis of the structure of arguments in ‘real’ (as opposed to purely mathematical or theoretical) situations. Toulmin proposed a structure starting with a *claim*, which is supported by *data* in accordance with the justification afforded by a *warrant*. Warrants in turn are justified by their *backings*; may be subject to *rebuttals*; and are associated with *modal qualifiers* (such as ‘usually’, ‘sometimes’, ‘some’) which qualify the degree of certainty with which the claim can be asserted. Our classifications of warrantability ‘by nature’ and ‘by substance’ correspond in Toulmin’s terms to different types of backings for warrants.
2. In propositional logic *modus tollens* is the inference rule: ‘from the premisses “*P* implies *Q*” and “it is not the case that *Q*”, conclude “it is not the case that *P*”’.
3. It might be suggested that arguing on the basis of the Home Office report is tantamount to using the ‘argument from auditee claim’ scheme discussed earlier, since the Home Office is the audited body here; but if the Home Office commissioned the report from experts, then the ‘expert opinion’ scheme is appropriate. (In practice, critical questions for the two schemes will be similar.)

REFERENCES

- Bechberger, E, D Lane, T McBride, A Morton, D Quintas and C Wong (2011), ‘The National Audit Office uses OR to assess the value for money of public services’, *Interfaces*, forthcoming.
- Bex, F, H Prakken, C Reed and D Walton (2003), ‘Towards a formal account of reasoning about evidence: argumentation schemes and generalisations’, *Artificial Intelligence and Law*, **11**, 125–65.

- Keen, J (1999), 'On the nature of audit judgements: the case of value for money studies', *Public Administration*, **77** (3), 509–25.
- Lonsdale, J (1999), 'The auditor's craft', in C. Pollitt (ed.), *Performance or Compliance? Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- National Audit Office (2004), *Managing Attendance in the Department for Work and Pensions*, London: National Audit Office.
- National Audit Office (2005), *Reducing Vehicle Crime*, London: National Audit Office.
- Pollock, J (1987), 'Defeasible reasoning', *Cognitive Science*, **11**, 481–518.
- Pollock, J (1995), *Cognitive Carpentry: A Blueprint for How to Build a Person*, Cambridge, MA: MIT Press.
- Pollock, J (2002), 'Defeasible reasoning with variable degrees of justification', *Artificial Intelligence*, **133**, 233–82.
- Prakken, H (2004a), 'Analysing reasoning about evidence with formal models of argumentation', *Law, Probability and Risk*, **3** (1), 33–50.
- Prakken, H (2004b), 'AI & law, logic and argumentation schemes', Session on Artificial Intelligence and Law, First Congress of Compared Legal Cultures and Systems, Instituto de Investigaciones Jurídicas, Universidad Nacional Autónoma de México, Mexico City, Mexico, 9–14/2 2004.
- Prakken, H and GAW Vreeswijk (2002), 'Logical systems for defeasible argumentation', in D Gabbay and F Guenther (eds), *Handbook of Philosophical Logic*, 2nd edition, vol. 4, pp. 219–318, Dordrecht: Kluwer Academic Publishers.
- Scharaschkin, A (2010), 'What is a conclusion on value-for-money?', Internal paper, National Audit Office.
- Schum, David (1994), *The Evidential Foundations of Probabilistic Reasoning*, New York: Wiley.
- Shafer, Glenn (1976), *A Mathematical Theory of Evidence*, Princeton: Princeton University Press.
- Toulmin, Stephen (1958), *The Uses of Argument*, Cambridge: Cambridge University Press.
- Verheij, B (2003), 'Dialectical argumentation with argumentation schemes: an approach to legal logic', *Artificial Intelligence and Law*, **11** (1–2), 167–95.
- Walton, D and C Reed (2003), 'Diagramming, argumentation schemes and critical questions', in FH van Eemeren et al. (eds), *Anyone Who Has a View: Theoretical Contributions to the Study of Argumentation*, pp. 195–211. Dordrecht: Kluwer Academic Publishers.

7. Forming judgements: how do VFM audit teams know what they know?

Justin Keen

Alex Scharaschkin's chapter examined sound argument construction in performance audit, drawing on work on the nature of evidence and argument in legal reasoning. This chapter seeks to complement that contribution by looking further at the question: how do audit teams employ data and methods in arriving at their judgements? The relevance of these arguments to the themes of the book is that how audit teams decide what it is they know will directly shape the evidence and argument they produce. This evidence and argument, in turn, shape the contribution of performance audit to accountability and democratic governance. This chapter presents a descriptive framework for discussing the strategies that performance audit teams use to produce their reports. In the first section, we consider the role of judgements in audit work before commenting on the literature on performance audit methods, and then presenting and discussing the framework.

JUDGEMENTS IN AUDIT WORK

In the last two decades and more, the practices of auditors have extended into new fields and, as we noted in Chapter 1, the term 'audit' has come to be used ever more widely (Power, 1997; Strathern, 2000; Lapsley and Lonsdale, 2010). One type of judgement is probably best known – the summative judgement or opinion provided by a financial auditor on sets of financial statements. This usually indicates that these financial statements are free from material mis-statement and are fairly represented in accordance with the required accounting regulations. Many are at the level of individual organisations, but judgements are also made on overall state accounts in many countries, and in the Statement of Assurance on the European Union Budget. In financial audit the judgements are typically presented as being 'true and fair' or as 'properly presenting' statements of accounts. If the audit team does not have sufficient evidence in a

particular area, or there is evidence of mis-statements, the judgement can be qualified.

In contrast, traditionally, performance audits have tended not to come to overall judgements. Because they have looked at a number of sub-issues or questions, they have usually focused on commenting on performance on each of these. Yet their role in accountability processes invites summative conclusions. Parliamentarians, senior government officials and other stakeholders – not least journalists – want to know what auditors think about an issue that they are reporting on. Who were the heroes, and who the villains? And how sure are the auditors in their judgements? This seems to have been recognised in the United Kingdom and, where possible, summative judgements are now included in the NAO's VFM reports in the form of a 'Value for Money conclusion', which are expected to answer (or at least address) the question – did this activity represent good value for money?

As Furubo noted in Chapter 2, much academic discussion about performance audit has centred on a debate about the relationship between audit and evaluation. Some influential commentators in both Europe and North America in the 1980s and 1990s – for example, Chelimsky, Leeuw, Pollitt and Summa – were particularly interested in the similarities and differences between performance audit and evaluation, and their respective roles in the evaluation of policy and programme implementation. They argued that performance audit has characteristics in common with evaluation and there were suggestions that the two were converging (see Wisler ed., 1996). The most widely quoted author is Eleanor Chelimsky (1985, 1990); indeed, her 1985 paper seems to have sparked this line of inquiry. In it, she argued that programme evaluations were of three types:

1. Descriptive: studies about the nature of a particular service or system.
2. Normative: comparison of actual performance with a reference standard.
3. Cause and effect: seeking to demonstrate how an event leads causally to others.

Chelimsky characterised performance audit as normative in nature and, therefore, more limited technically than programme evaluation. She went on to describe differences in approaches to research problems, suggesting, for example, that auditors tend by nature to think inductively, whereas programme evaluators think deductively. She concluded that there were fundamental differences between the two, but that the two sides had enough in common to learn from one another, particularly in the areas of study design and methods. Pollitt and Summa (1998) supported her

general position, arguing that the extent of any convergence in approaches and methods is limited by important differences in institutional arrangements for audit and evaluation. More recent work, for example by Gendron, Cooper and Townley (2007), and by Levitt, Solesbury and Ling (2006) provided further support for this position.

Keen (1999) approached the topic in a different way, focusing on the ways in which studies were designed, and arrived at rather different conclusions about the nature of performance audit work. Whilst Chelimsky and others focused on the overall characteristics of performance audit studies – their objectives, the evaluation criteria they used – Keen argued that the key to understanding them lay in the way they were built up from their constituent elements. They were best understood as sets of ‘building bricks’. Each study was composed of many distinct elements, which were arranged into a coherent pattern in a report. Performance audit teams separate their studies into discrete boxes of ‘micro issues’, with each box containing some data, a method for assessing the data, and a judgement. One important consequence was that teams used a range of data, and judgement criteria, in the course of any one study.

There may be straightforward reasons for this. The clearance process mentioned by Scharaschkin, by which the content of the draft report is discussed with those audited (Sharma, 2007), the visibility of the NAO’s VFM reports and their use as the basis for parliamentary accountability hearings mean that auditors and auditees are interested in each sentence in a published report. In the course of their work, therefore, audit teams have to bear in mind the confidence they will need to have in each statement in their report: each one has to be supported. This naturally leads them to focus on sequences of audit points. In the NAO study on ‘Supporting people with autism through childhood’ (National Audit Office, 2009), for example, the team identified 38 issues about current provision and performance which they wanted to explore by gathering and analysing data (Levitt, Martin, Nutley and Solesbury, 2010). This is not to say that they are uninterested in overall judgements (particularly now they are required to come to a ‘VFM conclusion’), just that they tend to build up their judgements in the manner of using bricks to build a house – the result being a plausible narrative, linking the observations together in a report – rather than by cross-matching separate findings. Conclusions therefore tend to aggregate a discrete set of findings, rather than synthesise evidence.

Hammond in his work on assessing human judgement and decision-making (1996) makes a useful distinction between two strategies for establishing the validity of any statement, or set of statements. The NAO has a strong preference for *correspondence* accounts, where teams seek

evidence that each of their judgements is supported by the available facts. This approach can be contrasted with judgements based on conceptually coherent accounts of events. In our everyday lives, we often make a judgement that an account of events ‘rings true’, even if there is not the evidence available to support it. This is an example of a *coherence* judgement, where we are, in effect, concluding that the account is coherent, in the sense that a series of statements fit together, even if we cannot verify that each one is true. Thus, a performance audit team might judge that the overall direction of travel in a programme merits a positive report, even if programme managers have made some avoidable mistakes along the way.

Sometimes these panoptic judgements appear in reports. For example, in 2010 the NAO published a report that examined NHS stroke services. The auditors wrote:

Improvements in acute care are not yet matched by progress in delivering more effective post-hospital support for stroke survivors, where there are barriers to joint working between the health service, social care and other services such as benefits and employment support. Patients and carers also lack good information about the services they may need and how to access them on discharge from hospital, as well as on how to prevent further strokes. (National Audit Office 2010, p. 8.)

Nevertheless, they concluded (p. 10) that:

The Department’s approach of developing a national stroke strategy underpinned by national and local leadership, national tier 1 performance indicators, clinical audit data, a national stroke tariff and £105 million seed corn funding, has increased the priority given to stroke care. Early indications are that implementation of the strategy is also starting to deliver improved levels of service and improved outcomes.

More often, however, reports are based on a series of correspondence judgements – the verification that each statement is true, and the creation of a narrative that links the ‘true facts’ into a readable report.

THE METHODOLOGY SQUARE

This line of argument can be extended, allowing more detailed consideration of the nature of performance-audit studies and similarities and differences with inspections and evaluations.

Figure 7.1 presents the Methodology Square. Running along the bottom of the square is a continuum linking two modes of thinking: intuition and analysis. Many authors have presented intuition and analysis as

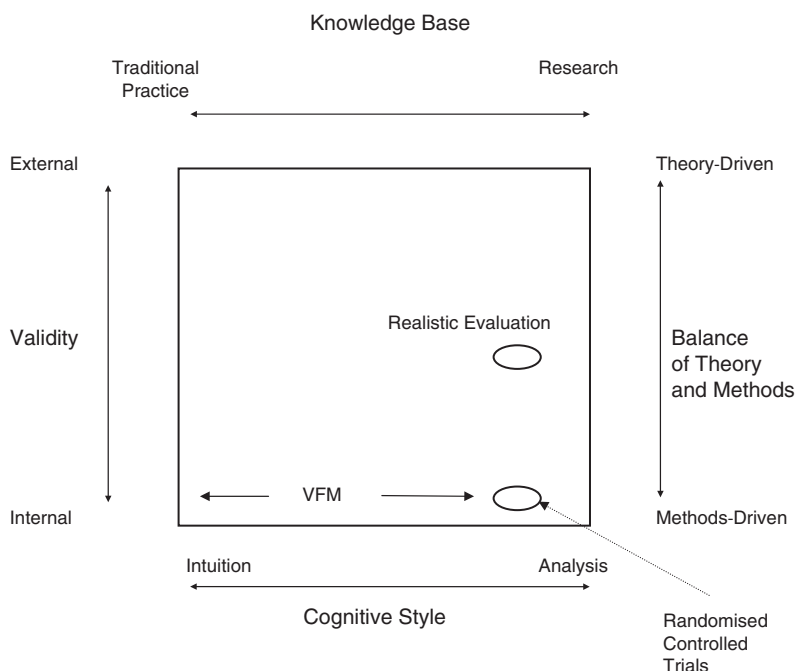


Figure 7.1 *The Methodology Square*

alternatives to one another, with intuition characterised by rapid, difficult-to-articulate thought processes, and analysis being structured, transparent and suited to situations where there is time and information available to structure one's thoughts. Hammond (1996) conceptualises the two as being linked by a continuum, however, and argues that many judgements we make are supported by mixes of intuition and analysis, and at different times we operate at different points on the continuum. This view is supported in the literature, for example, by Harper (2000) on the audit-style investigations of the International Monetary Fund.

The top edge of the square complements the bottom, with a continuum representing modes of practice. At one end is a position where practice is essentially intuitive in nature, for example, in the case of a lawyer or doctor arriving at a judgement on the basis of deeply embedded thought processes that he or she has difficulty in articulating: 'How do I know the patient has liver disease? I just know.' At the other end of the spectrum is research-based practice, where all decisions are based on empirical evidence and are transparent (in the sense that the logic of any decision

is explicit). This end of the spectrum may be aspired to by practitioners but often not reached, partly due to a paucity of evidence in many fields, and partly due to the problem of applying 'pure' research-based evidence directly to messy real world practice.

The left hand side of the square posits a continuum between internal and external validity of research findings. This is a novel representation but many discussions of the aims of evaluation discuss validity as if they have such a continuum in mind (Cronbach, 1982; Pawson and Tilley, 1997, Ch.1). Some practitioners perceive maximising internal validity as a goal, while others (notably Cronbach) advocate the development of approaches and designs that maximise external validity. This is not the place to discuss the merits of striving to maximise internal or external validity or the technical difficulties associated with approaching either end of the spectrum, save to note that advocates of more overtly scientific styles of evaluation have tended to lay stress on ways of maximising internal validity.

The right-hand side of the square represents the fourth continuum, this time between methods-driven and theory-driven research. Again, the continuum is implicit in the writings of a number of authors, including Pawson and Tilley (1997) on evaluation and George and Bennett (2005) in the context of case study methods. The continuum represents one of the main preoccupations of academic research and evaluation practitioners. Those who conduct randomised controlled trials, for example, are interested in methods and generally eschew theory (however it is defined). At the theory-driven end of the spectrum, one finds the theoretical frameworks that underpin disciplines, such as principal-agent or transaction cost theories in economics. More generally, those who believe that 'one size fits all organisations' reside here, including consultants who argue that all organisations can benefit, for example, from business process redesign or a particular new IT service.

Performance Audit and the Methodology Square

Where is performance audit located in the Square? If we consider performance audit reports as wholes (in the manner of Chelimsky and others) then we can say that they tend to emphasise internal rather than external validity, intuition rather than analysis, and methods rather than theoretical argument. They are therefore located towards the bottom left-hand corner of the square. If one looks at the constituent elements of reports though, the situation is more complicated; some elements are also properly located towards the bottom left-hand corner, but others range across the cognitive continuum between intuition and analysis, with parts of some studies being overtly analytical in nature.

These points can be illustrated by locating a value-for-money study in the square. In the NAO study of stroke services, there are statements about the geographical variation in payments found by the study team. For example, the auditors reported:

In April 2009, 71 per cent of hospital sites in England offered thrombolysis to stroke patients compared to 18 per cent in 2006. The total number of patients thrombolysed in 2008–09 was more than a two-fold increase compared to the previous year – with an estimated additional 1,000 patients receiving this treatment – and approximately five-times higher than in 2006–07. (National Audit Office, 2010, p. 24.)

The statement is based solely on empirical evidence and sits at the bottom right hand corner of the square. In an adjoining paragraph, the report also states:

In some regions, such as Greater Manchester and London, the ambulance service can take patients directly to designated hyper-acute stroke services, which can mean that stroke patients are not always taken to the nearest hospital. This approach may not be appropriate for all regions, such as those with geographically dispersed populations and hospitals. (National Audit Office 2010, p. 24.)

These sentences reflect more subjective judgements made by the study team. The team uses conditional statements, perhaps conscious that it is not possible to say that all areas of England will be able to adopt the same service model as Greater Manchester. A similar approach is adopted in relation to other quantitative data. Thus:

All hospital trusts in England now have a stroke unit, and the median number of beds per unit has increased from 24 in 2006 to 26 in 2009. The proportion of patients spending 90 per cent of their stay on a stroke unit (one of the Tier 1 'Vital Sign' measures) had increased from 47 per cent in the three months from January 2009 to 57 per cent six months later . . . [Comparison against national targets shows] that there will need to be significant improvements to reach the Department's expected position of 80 per cent by the end of 2010–11. (National Audit Office, 2010, p. 25.)

This statement is based partly on empirical foundations, but also partly on judgements about the extent of compliance with central guidance. It therefore reflects a mixed or compound judgement and, as with the last quotation, sits in the middle of the bottom edge of the square. Finally, as noted earlier, the report also makes overall judgements about stroke services, and these can be characterised as coherence judgements. These

are essentially intellectual judgements (albeit based on the amassing of a range of empirical evidence) and therefore sit in the middle of the Square.

Other Types of Evaluative Activity

Different types of evaluative activity can also be located in the square. In order to do this it is necessary to represent each style by its dominant approach and leave out any 'dissenters' within that style. This inevitably runs the risk of naïve generalisation, but the naïvete serves a purpose here. Proponents of randomised controlled trials, with their stress on analysis and internal validity, reside at the bottom right-hand corner of the square. Much programme evaluation resides close to, but not in, the bottom right-hand corner. Realistic Evaluation (Pawson and Tilley, 1997) seeks to move studies vertically up the square by incorporating theory and increasing external validity, while retaining the advantages conferred by rigorous empirical analysis.

It is important to stress what the Methodology Square does and does not tell us. It is intended to illustrate trends and tendencies: the dimensions represented were chosen because they provide both commonalities and differences between auditing and evaluation, and the choice was pragmatic rather than being based on any particular theory. Since the Square has no dimensions, the 'distance' between the different styles is not calibrated. The Square is essentially descriptive rather than normative in nature. It is also a matter of debate where one ought to be within the Square; there is no single best position.

Finally, it is worth noting that accountability and performance improvement studies are often presented as being in tension with one another (for example, Chelimsky, 1985; Lehtonen, 2005) but, as Pollitt and Summa (1998) and Barzelay (1997) note, the two are actually different types of study and imply different relationships between researcher and researched. The arguments presented here allow us to say what these differences are, at least in general terms. We can say that accountability studies of the kind undertaken by the NAO are based on detailed descriptions and occupy the bottom of the Methodology Square, whilst evaluations and case studies are concerned with providing explanations, with identifying change strategies or, in some styles of evaluation (for example, Patton, 1997) with securing tangible change.

The issue is not, therefore, whether or not performance audit supports learning rather than accountability, but what kind of evidence it can introduce to either of these activities. A concern with the tangible and demonstrable may trump an interest in intangible outcomes and hidden

mechanisms of change. This in itself is neither good nor bad, but it should help us to reflect on how performance audit adds value.

BUILDING VFM STUDIES

The Methodology Square allows us to re-interpret the literature on audit practices. Two arguments are pursued here. The first concerns the purposes of performance audit studies. The literature suggests that they can be exploratory, descriptive, explanatory or – more rarely – predictive. Looking at reports as a whole, as Barzelay and others argue, performance audit studies are typically descriptive. Some studies in little researched areas are essentially exploratory in nature. NAO studies can also contain predictions, albeit of a general nature. For example, the study of stroke services notes the financial pressures on public services and the difficulties that NHS organisations will face in coming years. So we can say that VFM studies tend to be descriptive in nature, but that their constituent elements can contain elements of exploration, explanation and prediction within them. These types of study can be mapped onto different purposes, as shown in Table 7.1. This undoubtedly simplifies the reality of both performance audit and evaluation, but highlights the point that different types of study may be suited to different purposes.

The Methodology Square allows us to take one more step. If the principal purpose of a study is to provide accountability, then it seems reasonable to argue that the auditor or evaluator should concentrate on empirical accuracy – on correspondence judgements – and hence stick to the bottom of the Square. If, as is often the case in evaluations, a key purpose is to explain a policy or programme, then one would expect there to be a role for theory, as the evaluators will be seeking generalisable – or externally valid – findings.

The second line of argument concerns the ways in which performance-audit studies are constructed. On the basis of the observations made about the NAO's work, it is possible to extend the analysis by looking at the construction process. Chelimsky (1985) argued that performance audit studies are planned and most subsequent authors have explicitly or implicitly agreed with her. But this chapter does not offer support for that view. Rather, the micro issues or elements are pieced together into a clear narrative: the NAO's reports are clearly written and can be read from cover to cover.

In short, VFM reports are constructed as narratives around each high level question. To return to the simple analogy used earlier, the process is somewhat like bricks being built up, so that the overall effect is clearly a wall. The essence of the approach lies in the fact that each brick has been

Table 7.1 Types and purposes of studies

Type of study	Purpose
Descriptive, Exploratory	Accountability
Explanatory, Predictive	Identifying potential change
Process Consulting, Action Research	Securing change

hand made; no two bricks are quite the same and the cement is provided by the narrative skill of the team. As noted earlier, the strength of the argument and evidence (in our analogy, the robustness of the wall) is also tested through clearance (the process of reaching consensus between auditors and auditees).

RELEVANCE TO THE AUDIT VERSUS EVALUATION DEBATE

In this author's view, these observations take us further away from the idea that performance audit – or at least the NAO's work – has much in common with evaluation. Performance audit and different styles of evaluation have fundamentally different characteristics, and yet may use some of the same methods (case studies, surveys, and so on). While the fine detail of these methods may vary, some of these methods do appear to be transportable between styles (many people use surveys and focus groups, for example) but the ways in which many other methods are used, while superficially similar, differ in important ways between styles. Interviews are a good example of where the VFM interview and qualitative research interview are quite different things. Many audit interviews are undertaken to enhance the auditors' understanding of the topic and for general illumination. They may pursue chance remarks made by interviewees if they judge them useful. Many qualitative research interviews are seeking specific information which will be used as evidence in the report, and teams may have detailed lists of questions to elicit specific, factual answers. What is more, the latter is not in any sense an evaluation method; it is simply a research method that happens to be used in basic research and in certain types of evaluation. It may therefore be helpful to think of performance audit and some styles of evaluation as drawing on a *common well* of ideas and practices, where those ideas and practices are drawn up into different institutional environments. Sometimes the method will be drawn up unscathed, but more often it will be transformed to the point where only the label ('interview', 'questionnaire') remains.

CONCLUSIONS

This chapter has focused on the nature of the technology of performance audit, primarily as operated in the NAO. It has argued that many performance audits can best be understood as a series of building bricks. An important consequence of this is that teams use a range of data and judgement criteria in the course of a study. The chapter has also drawn on wider thinking on human judgement to argue that audit teams tend to have a preference for correspondence, rather than coherence, judgements, where each of their statements are supported by available facts.

It has also outlined a framework for thinking about the nature of performance audit and the similarities and differences between audit and different styles of evaluation. It has argued that whilst overall, performance audit tends to emphasise internal rather than external validity, intuition rather than analysis, and methods rather than technical argument, individual elements of studies – because of the way they are constructed – can be placed at different spots on what we have depicted as continua on the Methodology Square. Thus, some statements made in audit reports are based solely on empirical evidence, whilst others are more analytical. Some contain a mix. While the framework does not lay claim to capturing all of the similarities and differences between performance audit and evaluation, it provides a simple, graphical device for discussing them.

REFERENCES

- Barzelay, M (1997), 'Central audit institutions and performance auditing: a comparative analysis of organizational strategies in the OECD', *Governance*, **10** (3), 235–60.
- Chelimsky, E (1985), 'Comparing and contrasting auditing and evaluation: some notes on their relationship', *Evaluation Review*, **9** (4), 483–503.
- Chelimsky, E (1990), 'Expanding GAO's capabilities in program evaluation', *The GAO Journal*, Winter/Spring 8, p. 51.
- Cronbach, L (1982), *Designing Evaluations of Educational and Social Programs*, San Francisco: Jossey-Bass.
- Gendron, Y, D Cooper, and B Townley (2007), 'The construction of auditing expertise in measuring government performance', *Accounting, Organizations and Society*, **32** (1–2), 101–29.
- George, A and A Bennett (2005), *Case Studies and Theory Development in the Social Sciences*, Cambridge MA: MIT Press.
- Hammond, Kenneth (1996), *Human Judgment and Social Policy*, New York: Oxford University Press.
- Harper, R (2000), 'International auditing: a case study of the IMF', in M Strathern (ed), *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*, London: Routledge, pp. 21–53.

- Keen, J (1999), 'On the nature of audit judgements: the case of value for money studies', *Public Administration*, **77** (2), 509–25.
- Lapsley, I and J Lonsdale (2010), 'The audit society: helping to develop or undermine trust in government?', in M Adler (ed) *Administrative Justice in Context*, Oxford: Hart Publishing, pp. 73–96.
- Lehtonen, M (2005), 'Accounting (f)or learning?', *Evaluation*, **11** (2), 169–88.
- Levitt, Martin, Sandra Nutley and William Solesbury (2010), *Evidence for accountability: Using Evidence in the Audit, Inspection and Scrutiny of UK Government*, London: Nuffield Foundation.
- Levitt, Ruth, William Solesbury and Tom Ling (2006), 'Evidence for Accountability. The nature and uses of evidence in the audit, inspection and scrutiny functions of Government in the UK', ESRC UK Centre for Evidence Based Policy and Practice Working Paper 24.
- National Audit Office (2009), *Supporting People with Autism through Adulthood*. HC556, Session 2008–09, London: TSO.
- National Audit Office (2010), *Progress in Improving Stroke Care*, London: National Audit Office.
- Patton, Michael (1997), *Utilization-Focused Evaluation*, 3rd Edition, Newbury Park, CA: Sage.
- Pawson, Ray and Nick Tilley (1997), *Realistic Evaluation*, London: Sage.
- Pollitt, C and H Summa (1998), 'Evaluation and the work of Supreme Audit Institutions', in M C Kessler, P Lascoumes, M Setbon and J-C Thoenig (eds), *Evaluation des politiques publiques*, Paris: L'Harmattan, pp. 63–82.
- Power, M (1997), *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press.
- Sharma, N (2007), 'Interactions and interrogations: negotiating and performing value for money reports', *Financial Accountability and Management*, **23** (3), 289–311.
- Strathern, M (ed) (2000), *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*, London: Routledge, pp. 21–53.
- Wisler, C (ed) (1996), 'Evaluation and auditing: prospects for convergence', *New Directions for Evaluation*, 71, Fall.

8. Standards and quality

Peter Wilkins and Richard Boyle

The quality of performance audits is crucial to the standing and credibility of Supreme Audit Institutions (SAIs). In the widest sense, quality in this regard relates to the effectiveness and efficiency of the performance audit programme of an SAI. While this is often considered through reviews of the overall performance of SAIs, much of the effort and interest in quality is focused on the processes that contribute to quality, as well as the quality of individual performance audits.

Lonsdale and Mayne (2005) detail two main aspects to quality in performance audit work, namely ‘the quality of the audit process (the ways in which reports are prepared) . . . and the quality of audit products themselves (the outputs from the work)’. The key characteristics of sound processes include independence, fairness and objectivity, and well-substantiated and careful production from inception to publication. Characteristics of quality in individual reports include accuracy, sound conclusions and recommendations, clarity, timeliness and impact. While this chapter addresses the concept of quality in the broadest sense, because of the focus on linkages between standards and quality, the main emphasis is on the quality of individual reports. Over the years, auditors have developed their own standards and created quality assurance arrangements to help maintain and improve quality. Mayne and Schwartz (2005, 9) note that ‘Professional standards are most evident in the area of performance audits and are strongly supported by audit offices as a means of ensuring credibility of their offices’. They also identify issues surrounding the setting of standards, including the authority behind the standards, and enforcement.

SAIs typically state that their performance audits are independent, impartial and unbiased and that the reports are evidence-based. They present their work as being authoritative. There are cases where the findings and recommendations of a report are challenged by the agency subject to audit and/or by stakeholders and experts, but these are relatively rare and usually occur in a context where the SAI is assumed by many others to have ‘got it right’. The chapters in this book by Scharaschkin and Keen indicate that performance audits produce contestable truths given the

nature of evidence and analysis involved, whilst the chapter by Leeuw sets out a range of criticisms of performance audits. The subject of standards and quality is therefore one of significance in defending the value of the work. This chapter therefore focuses on the differing approaches to performance audit standards and quality assurance in a number of countries, to draw lessons for wider consideration.

STANDARDS

Standards specify qualities or processes that are requirements. Bemelmans-Videc (2005) identifies standards as written instructions of a concrete nature, distinguishing them from guidelines which are of a more generic nature. The reasons for the creation of standards for performance auditing are not clearly specified but appear to relate to achieving consistency and at least meeting minimum standards of quality. Standards can also be an important tool for internal communication and aligning effort with the organisation's goals. The existence of clear standards can also help in communications with contractors and researchers who undertake or contribute to performance audits on behalf of SAIs. Compliance with a standard should give a user of a performance audit report assurance about the quality of the report. It should also inform the agency subject to audit about the appropriateness of the processes involved.

The absence of standards within an organisation creates the risks of inconsistent approaches, and this may in turn limit the ability to learn and improve by comparing across projects. Lack of consistent standards across organisations and countries carries similar risks and limitations. Conversely, standards may achieve consistency at the expense of innovation and adapting performance audits to specific contexts, the more so where the standards are of a prescriptive nature.

International performance audit standards, ISSAI 3000, have been set by INTOSAI as non-mandatory guidelines (INTOSAI, 2004) as part of a wider role issuing statements of Fundamental Auditing Principles and Auditing Guidelines. The guidelines emphasise that performance audits are wide-ranging analyses open to judgements and interpretations, and avoid elements that could be interpreted as conceptualising a single approach or prescriptive requirements (Table 8.1). INTOSAI (2004) notes that 'Advanced performance auditing is complex investigatory work that requires flexibility, imagination and high levels of analytical skills. Streamlined procedures, methods and standards may in fact hamper the functioning and progress of performance auditing. Consequently, standards – as well as quality systems – that are too detailed should be

Table 8.1 Comparison of performance audit standards

Characteristic	INTOSAI	NAO UK	GAO US	ANAO Australia
Basis	Guidance	Own motion	Own motion	Legislation
Mandatory	No	Yes	Yes	Yes
Source	Internal processes	Internal processes	Internal processes	AUASB
Objective	Independent, wide-ranging analyses of the economy, efficiency and effectiveness of government programs and agencies made on a non-recurring basis. Should have a focus on problems observed and possible causes	To report to parliament on the way in which government bodies use the resources voted to them, and to conclude on the value for money achieved	Engagements that provide assurance or conclusions based on a valuation of sufficient, appropriate evidence against stated criteria	To express a conclusion designed to enhance the degree of confidence of the intended users by reporting on the economy, efficiency or effectiveness of an activity against identified criteria
General principles applied	Wide-ranging and open to judgements and interpretations. SAI must have at its disposal a wide selection of investigative and evaluative methods	Competence Integrity Rigour Objectivity and independence Accountability Adding value Perseverance Clear communication	Independence Professional judgement Competence Quality control and assurance	Integrity, objectivity, professional competence and due care, confidentiality and professional behaviour

avoided.’ It also clearly differentiates performance auditing from financial auditing by stating that ‘while financial auditing tends to apply relatively fixed standards, performance auditing is more flexible in its choice of subjects, audit objects, methods, and opinions’ and emphasises that it is not linked with formalised opinions (page 12). Appendices to this standard give detailed guidance regarding matters such as audit criteria, evidence, documentation and quality assurance, and these are read as helpful guidance rather than prescriptive requirements.

INTOSAI gives a specific focus to the quality of performance audit reporting, specifying that the report should be reliable, complete, accurate, objective, convincing and as clear and as concise as the subject matter

permits. Unusually, it emphasises that this should ‘make the results less susceptible to misunderstanding’ and that ‘It should be easy for the reader to understand the purposes of the audit and to properly interpret the results’. The United Kingdom NAO’s Value for Money (VFM) handbook takes a similar tone (2009a page 3). It notes that ‘VFM work does not suit a ‘one-size-fits-all’ approach . . . however it is also important to recognise that we measure our professionalism against a set of principles which inform all of our VFM work.’ (page 4)

The NAO (2009b) also reports that in 2008–09 it established 10 explicit standards for its value-for-money work in its audit approach. These cover aspects of the work such as objective analysis, persuasive reporting, project management, evidence reliability and documentation, client engagement and report delivery. The NAO’s approach to its VFM work also for more than 10 years used five ‘quality thresholds’ – a set of questions that teams should ask themselves at different stages of the audit to enable them to decide whether they have met the standards expected. These are (page 29):

1. Are you ready to proceed?
2. Have you derived clear, compelling messages, supported by the evidence?
3. Are you ready to clear the draft report?
4. Are you ready to spread the message?
5. Have you learnt and disseminated the lessons?

Each question is accompanied by several sub-questions and sets out the role of key decision-makers. The approach is designed to recognise that VFM audit can be very varied and that it is better to encourage teams to assess themselves.

The United States Government Accountability Office (GAO) sets out its own performance audit standards which represent ‘generally accepted government auditing standards’ or GAGAS (GAO, 2007). Formally, under the standards, performance audits appear to be narrowly defined in terms of financial auditing concepts, being described as engagements that provide assurance or conclusions based on evaluation of sufficient, appropriate evidence against stated criteria. However, the standards for performance auditing are substantially broadened by noting that auditors have the option of using other professional standards in conjunction with GAGAS.

The Australian National Audit Office (ANAO) conducts its performance audits to the ANAO Auditing Standards which incorporate the Australian Auditing Standards (ANAO, 2009a).¹ The ANAO Auditing Standards make specific reference to ASAE 3500 regarding performance

engagements formulated by the Australian Auditing and Assurance Standards Board (AUASB) and indicate that it and other AUASB standards are incorporated 'to the extent that they are not inconsistent with requirements of the Act or other relevant legislation', the Act being a reference to the Auditor General Act 1997.

This AUASB standard has its origins in international standards as it is a subset of a generic standard for assurance engagements (ASAE 3000) which mirrors a corresponding international standard ISAE 3000, issued by the International Auditing and Assurance Standards Board of the International Federation of Accountants. However, Australia appears to be unique in having extended the generic international standard into a version tailored to performance audits. The concepts in the generic international standard have direct parallels with some of the GAO performance auditing approaches but are substantially different from those in the INTOSAI and NAO standards. The AUASB specifically narrows the purpose of a performance audit to expressing a conclusion, which is very different from the broad mandate provided by the legislation governing the ANAO. An additional implication is that this standard applies to the conclusion and not to the remaining contents of the report. The origins of this particularly narrow focus are the financial audit origins and focus on the issuing of an 'opinion'.

Unlike the other standards in Table 8.1, ASAE 3500 is intentionally written to be used by organisations other than audit offices, including for example, internal auditors or private consultants. It raises a different perspective on the independence of the performance auditor as it enables the auditor to be engaged by the management of the agency, with reliance based on the obligation in the standard for the auditor to behave independently. It also extends the scope relative to the remaining standards as it enables the auditing of private-sector entities.

By way of contrast, it is notable that the combined effect of the NAO's approach of using principles and quality thresholds is a much softer approach to standard-setting than the more prescriptive requirements of standards that draw their origin from financial auditing standards such as those issued by the AUASB. Similarly, the NAO's approach more clearly encompasses the research and methodological approaches of diverse professions and disciplines.

In summary, there is considerable variation between SAIs of the standards that apply to the conduct of performance audits. As indicated by the comparison across four institutions in Table 8.1, this in part flows from differences in the legislation establishing the SAI, but is also influenced by accounting-based standard-setting processes (with the influence evident in the standards of the GAO and ANAO).

In some ways, this difference in approach to standards across SAIs might be seen as unexpected. Given the degree of contact between SAIs and the exchange of experience that takes place between offices, some 'standardisation of standards' might be expected. However, in reality this variation is understandable. While SAIs are expected to meet high technical standards (which one would expect to be similar in each country) they also seek to be regarded as legitimate within the particular institutional and cultural context in which they operate (which would lead one to expect them to be different). What counts as evidence, how it is weighed, what are reasonable grounds of contestation and structures of argumentation are all in part determined by culture, context and history rather than epistemology, ontology and logic. So differences in standards and in their application across SAIs are to be expected.

The significance of the differences between the standards is illustrated here by examining their approaches to the levels of assurance/evidence and to recommendations. Some of these differences are highlighted in the following sections.

APPROACH OF STANDARDS TO ASSURANCE AND EVIDENCE

INTOSAI observes that an audit and the SAI conducting it must be independent, possess required competence and exercise due care. Evidence should be sufficient (to support the audit finding), competent (consistent with facts) and relevant (having a logical, sensible relationship to the finding). It also points to the importance of defining the study design by assessing the kind of information needed to answer the questions posed. It then distinguishes studies focusing on outcomes, processes, impact, cost benefit, benchmarking and meta-evaluation and discusses the types of study-design involved. The UK NAO adopts a similar approach, prefacing the section of its guidance on 'sufficient, relevant and reliable evidence' by discussing appropriate methodologies and illustrating typical methodologies used on VFM studies. It notes that 'A good VFM examination will probably include at least three or four different methodologies. Capturing a range of data and triangulating findings from different sources is an important way of building strength into the final report . . .'.

In contrast, the GAO adopts financial auditing concepts, noting that a performance audit provides reasonable assurance that evidence is sufficient and appropriate to support the findings and conclusions. Auditors must obtain sufficient, appropriate evidence to provide a reasonable basis for their findings and conclusions. The AUASB takes this interrelationship

with financial auditing concepts further with a performance audit potentially providing two levels of assurance. These are: reasonable/limited assurance where the objective is a reduction in performance engagement risk to an acceptably low level for a positive form of expression of the conclusion; or to a level that is acceptable in the circumstances for a negative form of expression of the conclusion. Understanding and applying this concept poses substantial intellectual challenges when translated to the complex and ambiguous performance issues explicitly recognised by INTOSAI and the NAO. The alternative approach, implicit in the softer standards, is to use precise language to convey the degree of assurance regarding any particular statement, with a continuum rather than a two-state model underpinning the intellectual approach involved.

While past practice in many SAIs would have seen individual teams selecting methods that worked well for them in the past, there are clear moves in the number of SAIs to train and equip their staff to be able to scan the range of methods available and select those which are appropriate to the purpose. Some standards more clearly address this approach than others.

APPROACH OF STANDARDS TO RECOMMENDATIONS

While all the standards reviewed make reference to performance audits possibly leading to recommendations and emphasise the important role they can play, they provide remarkably little guidance and certainly very few standards against which the quality of recommendations can be assessed.

INTOSAI recognises recommendations as a possible component of a report, noting that the audit should 'estimate the likely impact of the recommendations wherever possible'. It identifies them as courses of action suggested by the auditor relating to the audit objectives and based on the cause of a finding. It notes that they should be practicable and add value, and should be argued in a logical, knowledge-based and rational fashion. NAO succinctly poses the question 'Are the recommendations evidence-based and clear, and will they add value?' It has an extensive guide to 'Writing smart recommendations' which includes tips on generating better recommendations, such as having recommendations which are meaningful, specific, clear and useful, and ensuring that they enable an assessment of whether or not they have been implemented (NAO, 2008). The GAO records that 'Auditors should recommend actions to correct problems identified during the audit and to improve programmes and operations when the potential for improvement is substantiated by the reported

findings and conclusions. . . [They should] flow on logically from the findings and conclusions, be directed at resolving the cause of identified problems, and clearly state the actions recommended.’ The AUASB recognises recommendations as a possible component of a report, but provides no guidance or requirements on their basis or content.

EVIDENCE ABOUT THE MERITS OF DIFFERENT APPROACHES

The analysis above highlights that three SAIs with well-established performance audit programmes along with the INTOSAI have adopted substantially different approaches in the standards they have adopted. There are substantial differences between the approaches to performance-auditing standards both at the international standard-setting level between national audit offices, as well as between international standards and national audit office standards. For instance, the relevant standards of the International Auditing and Assurance Standards Board (IAASB) of the International Federation of Accountants is highly prescriptive when compared with the INTOSAI standards, and thus a country like Australia, which has built its performance auditing standards around the IAASB standards, is on a very different path from SAIs which have developed their own standards loosely shaped by the INTOSAI standards. This apparently somewhat confusing situation, which in part suggests that the intellectual foundations for the setting of performance audit standards is weak, also reflects the reality that interventions are applied in different ways in different contexts (as discussed above). It also opens opportunities for research and learning as the application of the different approaches progresses. Standards are set based on professional judgement, primarily by those involved in performance auditing, and it appears that the background of these institutions and individuals is a determining factor in the approaches adopted.

There is agreement between INTOSAI and the IAASB to cooperate with each other in the development of International Standards on Auditing with convergence being progressed at this stage for financial audit (Azuma, 2008). Recent discussions by the Performance Audit Subcommittee of INTOSAI resulted in the conclusion that the IAASB concept of assurance engagements is ‘of limited general relevance to the performance auditing conducted by advanced SAIs’ (INTOSAI Professional Standards Committee, 2009). Conversely, the IAASB is currently considering amendments to ISAE 3000 that provide more detail about its approach to performance audit type work (IAASB, 2010). This

suggests that if there is to be convergence between the different types of standards it may take considerable time and be a test of the relative influence of different professions.

It is surprising that there is no evidence base on which to identify the relative merits of the different approaches. While it may be that different standards are appropriate in different jurisdictions, there would appear to be merit in having future decisions underpinned by evidence of which approaches are most appropriate in particular contexts.

Implications for Quality

As standards could both increase consistency of approach (and thereby facilitate institutional review) and limit options, it would seem important to understand the relative benefits and disadvantages involved. The overriding objective of performance audit standards is to maintain and strengthen the quality of the work undertaken, to ensure consistency and to give confidence to the users of the information. The next section of this chapter discusses how the quality of performance audits and performance-audit programs is assessed, although, interestingly, the work regarding quality has rarely been linked back to the types of standards involved.

As identified earlier, the adoption of the standards by an institution conveys the message that it is concerned about quality and consistency, and has made a commitment to achieve at least the minimum requirements contained in the standards. To the extent that the standard applies to other institutions, it shows that the SAI is committed to meeting at least the minimum requirements that apply to a wider community of performance auditors.

The adoption of externally generated standards that are voluntary raises questions about why this decision was made. What is the case for the adoption of external standards as the best way to improve effectiveness and efficiency? Might an independent auditor-general be seen to be subjugating his or her independence to some other standard-setting body? While in general SAIs accept that their financial audits will be conducted to externally set standards, there is at least one example of this being challenged. In a recent submission to a parliamentary enquiry in Australia, the Victorian Auditor-General called for a lifting of the legislated requirement that all his audits (including performance audits) be conducted according to the Australian Auditing Standards. He proposed that he have a limited discretion to dispense with an auditing standard, fully or partially. The Auditor-General stated 'As there are few standards written for the public sector, circumstances may arise where compliance with the standard may conflict with powers or duties under the Act . . .

In the light of this possibility, a means for the Auditor-General to opt out of an auditing standard becomes necessary in such circumstances'. In its Discussion Paper, the parliamentary committee specifically notes the role of ASAE 3500 in relation to performance audits and has sought comments on the issues raised by the Auditor-General (Victorian Public Accounts and Estimates Committee, 2010). Questions also arise as to whether the auditor has aligned his thinking with a particular profession or professional approach rather than working from a broad public interest perspective. These issues clearly become more pointed in jurisdictions similar to Australia and the US which have more prescriptive standards.

The adoption of the standards as mandatory requirements also creates a risk of non-compliance. Where a conscious decision is made not to comply and this is reported publicly, it raises questions in the reader's mind as to whether the non-compliance was warranted and what are the implications for the quality of the report. If any non-compliance is not fully reported publicly, later highlighting of this could undermine confidence in the report and potentially the institution. Where non-compliance is inadvertent there are similar risks of these being identified through internal and external quality reviews or by users more generally, again undermining confidence.

To the extent that the standards are prescriptive, they may impose process requirements that are not warranted on the basis of efficiency and cost-effectiveness. The implementation may also have a wider and more diffuse influence on quality by limiting the types of performance audits undertaken and restricting the culture of the organisation in relation to its performance audit function. For instance, performance auditing dominated by a single professional discipline may narrow the range of professional backgrounds and personal skills of staff attracted to work and stay in the organisation.

The cultural influences are difficult to identify, but are illustrated by the wording of the AUASB standard, the users of which need to be able to comprehend concepts and text such as the following which may be very alien to their professional and general experience (AUASB, 2008, 10):

The assurance practitioner shall evaluate, individually and in aggregate, whether:

- (a) deficiencies in systems or controls; and/or
- (b) variations of the measures or assertions from the identified criteria that have come to the attention of the assurance practitioner are material to the conclusions in the assurance report.

Use of language of this kind makes the standards for performance auditing accessible to those from a financial audit background but would tend

to exclude performance audit staff from other disciplines. Standards of this type represent a barrier to entry by other professions and pose challenges to the many performance audit units that are established as multi-disciplinary teams. Linked to a view of differing culture, is an analysis that suggests efforts to shape and control the professional basis of performance auditing. This is most evident in the IAASB standard ISAE 3000 that requires the work to be undertaken by public accountants and in relation to quality control make specific reference to ‘firms’ (ISAE 3000, 2005).

QUALITY

As with standards, definitions and interpretations of quality of performance audit are not simple or straightforward and interpretations can vary from country to country. Lonsdale and Mayne (2005) discuss what is meant by quality in a performance audit context. They note that the concept of quality varies over time, with greater attention in recent years to what outsiders think and the needs and expectations of users as opposed to more ‘traditional’ interpretations of quality as guaranteed by the status and authority of the SAI (page 178). They note two particular aspects of quality in performance-audit work: ‘. . . the quality of the audit process (the ways in which reports are prepared) . . . and the quality of audit products themselves (the outputs from the work)’ (page 179).

In general terms, many SAIs have well-developed quality assurance systems for financial audits, designed to provide assurance to the head of the office that appropriate procedures are implemented which ensure the quality of the individual audit assignments. These quality assurance systems are based on the requirements of relevant auditing standards. Notwithstanding the existence of standards or guidance regarding quality assurance in relation to the reporting of performance audits, SAIs have tended to develop and implement their own systems and procedures to try to ensure that performance audit outputs are produced to high quality standards.

INTOSAI (2004) specifies that a performance audit report should be reliable, complete, accurate, objective, convincing and as clear and as concise as the subject matter permits. As noted above, INTOSAI emphasises that reports should ‘make the results less susceptible to misunderstanding’ and that ‘It should be easy for the reader to understand the purposes of the audit and to properly interpret the results’. It distinguishes:

- Quality assurance – refers to policies, systems, and procedures established by the SAI to maintain a high standard of audit activity. A

quality assurance review program is identified as a series of external and internal reviews of activities undertaken by the SAI (p. 120).

- Quality control – refers to reviews conducted to assess the quality assurance system or the executed audit projects (p. 118).

SAIs typically try to assess their effectiveness using a range of performance indicators including the satisfaction of members of parliament, percentage of recommendations accepted and implemented, savings and other benefits generated, and so on. In some jurisdictions audit offices are subject to external reviews under their legislation at regular intervals, typically with a parliamentary committee responsible for the initiation and reporting of the review. Generally, these result in a value-for-money study of (amongst other things) the value-for-money function of the audit office. In the broadest sense these measures give some indication of whether the performance audit work is ‘fit for purpose’, with the choice of measures relating to the purposes identified for the individual SAI.

In addition to this, in some jurisdictions there is a structured process of peer reviews where peers from other audit offices conduct a value-for-money study of the SAI. A variant of this is having a peer review specifically focused on the performance audit function. Typically a team of reviewers, drawn from senior staff in the performance audit groups in one or two audit offices, conducts a review that may be as broad as the performance-audit component of the parliamentary review or something narrower.

Within audit offices there are typically many quality assurance processes, not only through line-management supervision, but also including off-line staff and senior groups of staff reviewing performance-audit projects at specific stages of their progress. Some audit offices structure their approach to quality assurance around the International Standards Organisation quality management systems requirements, with the New South Wales Audit Office first achieving ISO 9001 certification in 1996 (New South Wales Audit Office, 2010).

A more narrow focus on quality relates to review of individual performance audit reports. The quality of individual reports can be assessed by external review of the reports themselves after they have been completed. A study of practice by Lonsdale and Mayne (2005) identifies five reasons why the quality of information produced in performance audit reports is or should be important to SAIs. These are:

- SAIs are public bodies with a statutory role and it is reasonable to expect such bodies to perform to the highest standards.

- As public bodies with statutory powers to ensure change is considered by government, there should be expectations of quality. Encouraging inappropriate or ill-founded decisions could be wasteful and potentially dangerous.
- Those who promote quality need to demonstrate it themselves.
- The status and authority of SAIs depend on the credibility of their work.
- Given the use of public money, it is reasonable that the efficiency and value of an SAI should be assured, albeit in ways that do not threaten the independence of the audit body.

PEER REVIEW OF PERFORMANCE AUDITING

Peer reviews of audit practice are increasingly common in the international auditing community and a means of providing assurance about both the quality of the performance audit process and the performance audit outputs. The review is an evaluation of the performance of an SAI, generally by members of one or more other SAIs. Because of the necessary independence of an SAI, there may be few mechanisms for oversight of its activities within its own jurisdiction, and review by its international peers is intended to compensate for this. With regard to the performance audit function, the peer review might include a detailed review of a sample of recent reports. This may incorporate a review of the working papers and other information on the processes generating individual reports (including potentially a consideration of issues such as adequacy of evidence, rigour in data collection and analysis and effectiveness of communication with stakeholders over the life of the project). Some quality review work may also look at management processes in and around the organisational unit producing the reports.

ANAO (2009b) indicates that it is leading an international peer review of the Office of the Auditor General of Canada to provide an independent opinion on whether its quality management system is suitably designed and operates effectively and that it includes coverage of the performance-audit work (pages 6, 28). It comments that it also completed a review of two New Zealand Audits Office performance audits and participated in a review of a small number of performance audits undertaken by the Victorian Auditor-General's Office. The case study discussed in Box 8.1 provides more information about a peer review.

Often the peer review reports are internal documents. The Irish report, as noted in the case study, was published. So too was the peer review of the Netherlands Court of Audit (NCA), which has publicly released the

BOX 8.1 PEER REVIEW OF PERFORMANCE AUDITING: CASE STUDY

A peer review of the Office of the Comptroller and Auditor General in Ireland was undertaken in 2008. Included in this review was an assessment of the performance of the Office in carrying out performance audits (termed value-for-money audits in Ireland). Two studies were selected for consideration. The reviewers spoke to representatives of each auditee, and to the lead auditor for each study. They also reviewed the Office's Manual on Reporting which provides guidance on all the Auditor General's reporting outputs, and received general comments on the value-for-money output from other interviewees. Given the high level nature of the assessment, the reviewers did not examine the files for the two studies. Nor did they analyse the two reports. Instead, the focus was on the efficiency of the examination process and the perceived effectiveness of the work carried out.

In conducting the review, the peer reviewers chose to focus on four aspects which they saw as important from a quality and developmental perspective:

- *Topic selection.* The reviewers note that it is important that the subjects selected for review are of value to the parliament and other stakeholders and, above all, to the entity subject to the study. The determination of criteria for the selection and prioritisation of topics for performance audit was stressed.
- *Engagement with the value-for-money client.* The reviewers note that the general focus of value-for-money auditing, while supporting the accountability of public entities, should be performance improvement. They state that this is best achieved by auditor and client working with, rather than against, one another. In this context it is seen as essential that there be good engagement between auditor and client.
- *Methodology.* The reviewers focused on three issues here:
 - Engagement of expert advice by the audit team – the review notes that performance auditors are likely to be generalists and that the involvement of a subject expert can enhance the quality of the audit.

- Handling a change of scope – new matters may arise which may need to be brought within the scope of the study.
- Use of standards – the review notes that the standards and guidelines in use in the Office ‘take account’ of the INTOSAI Implementation Guidelines for Performance Auditing (July 2004) and of the INTOSAI Auditing Standards. They note, however, that the value-for-money methodology has not been made compliant with International Auditing Standards and see merit in the application of those standards to the value-for-money audits, particularly the more comprehensive quality assurance processes which the standards contain.
- *Staff skills*. The review notes that value-for-money auditing requires people with good analytical skills, and the ability to write the reports.

Interestingly in the context of this chapter, the Comptroller and Auditor General accepted most of the recommendations made by the peer reviewers, but with regard to the issue of standards raised by the review team he only partially accepted the recommendation and notes: ‘The Office’s manual on reporting, which covers value for money examination methodology, was developed based on the INTOSAI standards for performance audits. The extent to which International Auditing Standards contain additional standards of relevance to value for money appears limited’.

Further information available at: Office of the Comptroller and Auditor General (2008, pp. 27–32).

report of a peer review conducted by staff of the Norway, New Zealand, South Africa and United Kingdom SAIs (Netherlands Court of Audit, 2007). The terms of reference of the peer review focused on the quality of the performance audit reports and their impact, drawing on six performance audits that were published in 2005–06 and on impact assessments. It examined the NCA’s policies and carried out extensive interviews with people within the NCA, members of parliament, civil servants and other relevant stakeholders. The review did not examine the working papers of the audits. The report provided assurance that ‘performance audit reports from NCA are in accordance with good professional practice of

performance auditing, and provide Parliament with objective and reliable information on government performance’.

The NCA report records a 2003 decision by a group of 16 SAIs to perform peer reviews of the work carried out in the respective organisations. Interestingly, in the context of this chapter, there is no specific reference to criteria or standards applied during the peer review. An exception is the commentary on recommendations, where the comments make reference to the NCA’s WWWW (Why Who What How When) criteria for good recommendations. The report noted that some people interviewed preferred rather general recommendations, whilst others preferred more specific recommendations, and found that ‘overall the analysis of the audit showed that the recommendations are not fully meeting NCA’s own WWWW criteria’.

Peer review may also be used to complement and oversee internal quality assurance procedures. The Western Australian Auditor General (Office of the Auditor General Western Australia, 2009) notes that he has adopted the Australasian Council of Auditors-General (ACAG) Self Assessment and External Review Framework as the primary means to identify, measure and compare improvement opportunities. It explains that ‘this framework was developed from a range of assessment systems. It includes a conceptual framework supported by sets of graded standards, based on client expectations and legislative requirements, against which to measure performance. It embraces not only audit functions and divisions, but all activities of the business, including corporate services.’ He indicated that he also intended that during 2009–10 the Office would be assessed against the QA framework through external peer review conducted and coordinated by ACAG.

EXTERNAL REVIEW OF PERFORMANCE AUDITING

In some jurisdictions an external review of the SAI is required by legislation. The legislation which establishes the Australian National Audit Office empowers an independent auditor to conduct a performance audit of the ANAO at any time. The potential scope of such audits is very broad and the same as that which the ANAO might apply to a performance audit of another public sector agency. The most recent wide-ranging performance audit of the ANAO specifically looked at its performance audits but did not make any findings or comments regarding compliance with the specific standards that apply (Commonwealth of Australia, 2004). The reviewer noted that his own audit ‘has been performed in accordance with

Australian Auditing Standards applicable to performance audits . . . and accordingly included such tests and other procedures that are considered necessary in the circumstances’.

SELF-INITIATED QUALITY REVIEW OF REPORTS

Attention has also been given to initiatives such as external quality review of individual reports, as part of a broader quality assurance system, to improve the credibility and use of the reports. The NAO in the UK has been commissioning quality reviews of their VFM reports since 1993. A team of academic specialists examine and comment on the reports. External review has been found to be a useful stimulus to thinking about new methods, and as a direct incentive to improve (Lonsdale and Mayne, 2005; Lonsdale, 2008, p. 245). In 2004, the NAO moved to external quality reviews at draft report stage, rather than post-publication, although at the time of writing it had decided to move back to post-publication. The NAO (2009b) has also introduced internal ‘cold reviews’, designed to assess – on a sample basis – adherence by teams to the expectations set out in the VFM Standards introduced in 2008. The approach to external quality review taken by the Irish Comptroller and Auditor General is set out in Box 8.2.

OPTIONS FOR QUALITY REVIEW

The preceding discussion clearly shows that quality review is being taken increasingly seriously by SAIs. A desire to illustrate that the SAI adheres to good-practice quality assurance procedures and processes, and produces high-quality products is in part being driven by a more general trend in the public service to publicly prove the quality of services provided for the resources allocated to the function. But quality review can impose costs and burdens of its own, and it is important that SAIs, as with all organisations, run quality assurance processes that add value at reasonable cost. With this in mind, there are a number of options that SAIs need to consider when undertaking quality review. For example, to what extent should quality review be conducted internally or using external assessors? From the above discussion, it can be seen that there are a broad range of options available to SAIs considering quality review of their performance audit function. Internally, hot and cold reviews of performance audit reports by staff of the office and assessment of the quality of outcomes achieved by such means as measuring client satisfaction, percentage of recommendations accepted, savings made, and so on, can help give a picture of quality.

BOX 8.2 EXTERNAL QUALITY REVIEW OF PERFORMANCE AUDIT REPORTS – CASE STUDY

In Ireland, the Office of the Comptroller and Auditor General commissioned a quality review of a sample of performance audits in 2008. In drawing up the main and sub-criteria and scoring system to assess quality, various national and international sources were drawn on.

Seven main criteria were agreed, against which the VFM examination reports were assessed. These were:

- The background, context, scope and purpose of the report are clearly set out.
- The report is well structured and presented.
- The summary of findings puts across the main conclusions.
- The methodology is clearly set out and appropriate.
- The report is fair and balanced.
- The report's findings and conclusions are informative and based on the evidence and analysis of the data.
- The conclusions clearly show performance outcomes and recommend appropriate actions to improve VFM.

A five point scoring range was agreed, ranging from five for excellent to one for poor. The five point scoring range was found to be sufficient for assessing the quality of reports, and provides the opportunity for some international benchmarking given its use in other contexts (such as the UK National Audit Office VFM reports and some EU evaluation reports).

After agreement with the Office on the criteria and scoring system the review team met to ensure that all reviewers had a common understanding of the meaning of the various criteria and sub-criteria, the workings of the scoring system, and the approach to be adopted. At this stage, the seven individual reports were allocated to competent reviewers, scored against the agreed criteria, and a short review produced for each report. With a view to maintaining a consistent approach, each review and report was checked by a second reviewer and the team leader. This second reviewer function focused on ensuring that consistent standards were being applied to the reports.

Following on from the review of the individual reports, an overall assessment of the quality of the VFM reports was made. This overall assessment focused on two main issues:

- The drawing out of common lessons learned from the review of the individual reports with regard to areas where there is scope for improvement in the production of VFM reports by the Office.
- Highlighting particular examples of good practice that may be of assistance in the production of VFM reports by the Office.

However, such internal processes need to be complemented by external quality review. This is not least because increasingly, public service organisations are expected to provide independent verification of the quality of their performance reporting. SAIs are not exempt from this trend. Indeed, as we suggested at the start of this chapter, there is an onus on SAIs to model good practice: to lead by doing rather than lead by telling.

With regard to external quality review, there are possibilities to review (a) the overall systems, processes and outcomes achieved by the performance audit function of the SAI, and (b) the performance audit reports produced. In the case of the former, an overall scrutiny of the performance audit function by peer review appears to be the most frequently applied quality review option. An issue here is how to ensure consistency of interpretation. As was noted above, there are clearer, more widely accepted international standards for assessing financial audits than there are for assessing performance audit. Issues that need to be considered are the independence of the reviewers and the scope of the assessment, with the options for the latter ranging from issues of compliance with standards through to a full assessment of effectiveness and efficiency. In this context, it is important that the status of the reviewers and the standards and criteria used in any review process of the performance audit function are made explicit.

With regard to external quality review of performance audit reports, Table 8.2 sets out three main options and the following text explores the strengths and limitations of each option.

Option 1: Do Nothing, Relying on Internal Management and Supervision Alone

As regulatory impact-analysis practice suggests, it is useful to include a ‘do nothing’ or no policy change option as a benchmark against which to

Table 8.2 Options for external quality review of performance audit reports

	Advantages	Disadvantages
Option 1 Do nothing, rely on internal management and supervision alone	<ul style="list-style-type: none"> ● No additional costs ● No additional procedures for performance auditors to follow 	<ul style="list-style-type: none"> ● Does not provide external assurance of quality ● May lose out on different perspective on quality issues
Option 2 A. <i>Ex-post</i> summative external review of a group of reports	<ul style="list-style-type: none"> ● Provides opportunity for a structured look at quality ● Delivers lessons to improve future performance audit practice 	<ul style="list-style-type: none"> ● Does not benefit the reports reviewed
B. <i>Ex-post</i> summative external review of individual reports	<ul style="list-style-type: none"> ● Provides more timely information than Option 2A 	<ul style="list-style-type: none"> ● May lose overview lessons that review of a group of reports gives
Option 3 A. <i>Ex-ante</i> formative external review at final draft report stage	<ul style="list-style-type: none"> ● Enables performance audit report to be amended at draft stage and improves quality of the report 	<ul style="list-style-type: none"> ● May lose overview lessons that review of a group of reports gives
B. <i>Ex-ante</i> formative external review at business case and final draft report stages	<ul style="list-style-type: none"> ● Feedback of information at business case stage may short-circuit quality problems 	<ul style="list-style-type: none"> ● External reviewer may be more prone to be 'captured' by the process and reluctant to provide 'negative' views on draft report

assess other options. It may be that the added costs associated with introducing external quality review outweigh the benefits, and that it may be preferable to operate without the introduction of a formal external review process. The advantage of doing nothing is that there are no additional costs or procedures for performance-audit reports to go through. In this option internal quality review procedures operated by the SAI, such as normal line manager review, are seen as sufficient to provide assurance on the quality of reports produced.

A disadvantage of the 'do nothing' approach is that it fails to provide external assurance as to the quality and standard of work undertaken. External quality assurance may also, by tapping into expertise not readily

available to SAIs, prove useful in bringing a different perspective to bear on such issues as method choice, report presentation and so on.

Option 2: *Ex-post*, Summative External Review of Performance Audit Reports

In this option a group of completed reports are sent to external reviewers for quality review. This may occur on a regular basis, such as annually or bi-annually. The main advantage of this approach is that it provides an opportunity for a structured look at the quality of performance audit reports produced, and the identification of any general lessons that may emerge, in areas such as methodology, summary of findings and so on. The main benefit is to provide lessons to be learned for future performance audit work.

Disadvantages associated with this approach include the reciprocal of the benefit outlined above, namely that the exercise fails to be of benefit to the reports reviewed, as they have already been completed and gone through the audit process. It may also be a relatively costly financial option, depending on the number of reports to be reviewed. A reasonably large sample is needed to provide a comprehensive picture.

A variant of this approach is that an *ex-post* summative external review is carried out of each individual published performance audit report. This was the practice in the UK National Audit Office in its first phase of external quality review of performance audit reports in the 1990s and until 2004. Generic lessons may be lost in this approach although the reviewers provided an annual digest of common themes, and the comments on individual reports were only of value to future work since the report examined was already published.

Option 3: *Ex-ante*, Formative External Review of Individual Performance Audit Reports

The UK National Audit Office moved from *ex-post* review of individual reports (which they labelled ‘cold review’) to *ex-ante* assessment of individual draft reports (‘hot review’). They moved to this process as it enables comments made on the quality of the report to be fed into report finalisation, and the report amended accordingly if necessary. The main advantage is that quality review at the drafting stage enables the report to be amended in response to comments received and a higher quality report produced as a result.

A potential disadvantage of this option is that the focus on the individual report may mean that lessons learnt from examination of a group of reports fail to emerge. General trends or common issues emerging may not be

spotted or highlighted. It is also an additional 'burden' for the performance auditor to be completed during the course of an examination and, unless managed well, may have potential to delay the process (though there are no reports of this happening from the UK National Audit Office experience).

Within this *ex-ante* approach, there are two main sub-options for external review. One is to limit external review to comment at the stage where a draft report has been produced. The other option, also used for a while in the UK National Audit Office case, is to allow some external quality review input at the stage when terms of reference and the methodology are being determined for the examination. In this case, the advantage is that external expert guidance on methodological choices and drafting of evaluation questions can input to the process and short circuit problems that may otherwise arise later. A potential disadvantage is that the external reviewer becomes 'captured' by the process and more reluctant to give a 'negative' review at final draft stage. There is also a greater financial cost involved.

There are advantages and disadvantages associated with all the options outlined. And options two and three are not mutually exclusive. It is possible to envisage a quality review process that consists primarily of *ex-ante* review at the draft individual report stage, with periodic *ex-post* review of a group of reports to identify general lessons emerging.

CONCLUSION

It is not surprising that SAIs accord high priority to the quality of performance audits given their role in relation to accountability in democratic societies. It is fundamental to their standing and credibility and thus to the use made of the reports in accountability contexts. At a time when public trust in government institutions is a cause of concern in many countries, it is more important than ever that SAIs, as pillars of national accountability regimes, are not only providing, but are seen to be providing, independent and high-quality work.

It has been argued that standards and quality are inextricably linked, and that the best way to achieve quality is through setting and complying with standards. While there is an intuitive appeal to such statements the analysis provided in this chapter has identified that much is still to be learnt about what types of standards should be adopted, and the mix of quality review measures that will best serve the differing needs of different SAIs. There is no agreed set of standards that addresses all situations and that SAIs are applying in common.

These differing approaches to standards by different SAIs are to some extent explainable in terms of adaptation to local needs. However, the



significant differences between standards indicate the potential for very different futures for performance auditing. It is argued that a multidisciplinary approach opens the door to flexible and adaptable responses to changing contexts and circumstances to a greater extent than a single discipline approach derived from the sphere of financial audit.

The actual impact of differing types of standards depends on the strength of the purportedly 'inextricable link' with quality. The substantive role played by standards is called into question by the observation that most quality-review processes define their own criteria, rather than being directly linked to pre-existing standards established by the SAI. It could be that in many cases the adoption of standards is a symbolic act in support of maintaining the SAI's standing and credibility, and that there are very real practical challenges associated with translating global standards to fit local contexts.

Where compliance with the standards becomes a focus of SAI activity, it has the potential to increase consistency, although depending on the type of standards this may come at a cost in terms of limiting responsiveness and innovation, along with significant overhead costs.

While performance audit has well-established roles in democratic accountability, it is recognised by practitioners, supporters and critics alike that maintaining the quality of the work is central to the continuation of these roles. This chapter also argues that quality review processes provide tangible benefits to SAIs and their products, and through this pathway they also contribute directly to their standing and credibility. It is thus important that SAIs continue to explore the relative benefits of options such as relying on internal management and supervision alone, 'cold' summative external review of reports and 'hot' formative external review of individual reports. The pursuit of quality is an evolving process.

NOTE

1. It is a requirement of the legislation that the Auditor-General issue standards that have to be complied with by persons conducting performance audits under the legislation and the content of these standards is left to the discretion of the Auditor-General.

REFERENCES

- ANAO (2009a), 'Australian National Audit Office (ANAO) Auditing Standards', available at: [http://www.comlaw.gov.au/ComLaw/legislation/LegislativeInstrument1.nsf/0/10B56482A309E764CA2575F4000EAFCB/\\$file/ANAOAuditingStandardsAssignedbyAG25June20092.pdf](http://www.comlaw.gov.au/ComLaw/legislation/LegislativeInstrument1.nsf/0/10B56482A309E764CA2575F4000EAFCB/$file/ANAOAuditingStandardsAssignedbyAG25June20092.pdf) (accessed 31 December 2010).

- ANAO (2009b), 'The Auditor-General Annual Report 2008–2009', available at: www.anao.gov.au (accessed 10 March 2010).
- AUASB (Australian Auditing and Assurance Standards Board) (2008), *Standard on Assurance Engagements ASAE 3500: Performance Engagements*, available at: www.auasb.gov.au (accessed 10 March 2010).
- Azuma, N (2008), 'The framework of INTOSAI Government Auditing Standards: in the stream of international convergence', *Government Auditing Review* **15**, 77–97, available at: <http://www.jbaudit.go.jp/pr/pdf/e15d05.pdf> (accessed 10 March 2010).
- Bemelmans-Videc, Marie-Louise (2005), 'Guidelines and standards: assuring the quality of evaluation and audit practice by instruction', in Rob Schwartz and John Mayne (eds.), *Quality Matters: Seeking Confidence in Evaluation, Auditing and Performance Reporting*, New Brunswick, NJ: Transaction Publishers, pp. 145–72.
- Commonwealth of Australia (2004), *Report by the Independent Auditor of the Australian National Audits Office on the Results of a Performance Audit of 'Value for Money' provided by the ANAO*, available at: http://cms.anao.gov.au/uploads/documents/Performance_Audit_of_Value_for_Money_provided_by_the_ANAO_24-6-04.pdf (accessed 15 May 2010).
- IAASB (2010), *International Assurances Standards Board Background Papers for Agenda Item 4 – ISAE 3000 of the Meeting of March 2010*, available at: <http://www.ifac.org/IAASB/Meeting-BGPapers.php?MID=0207&ViewCat=1273> (accessed 15 May 2010).
- INTOSAI (2004), *International Organisation of Supreme Audit Institutions ISSAI 3000 Implementation Guidelines for Performance Auditing*, available at: [http://www.issai.org/media\(708,1033\)/ISSAI_3000E.pdf](http://www.issai.org/media(708,1033)/ISSAI_3000E.pdf) (accessed 15 May 2010).
- INTOSAI Professional Standards Committee (2009), *The Agenda for the IAASB Meeting in December 2009*, available at: [http://psc.rigsrevisionen.dk/media\(1180,1033\)/Executive_Briefing_November_2009.pdf](http://psc.rigsrevisionen.dk/media(1180,1033)/Executive_Briefing_November_2009.pdf) (accessed 15 May 2010).
- ISAE 3000 (2005), *International Standard on Assurance Engagements 3000: Assurance Engagements Other Than Audits or Reviews of Historical Financial Information*, International Federation of Accountants IAASB Assurance Engagements, available at: http://www.ifac.org/members/DownLoads/ISAE_3000.pdf (accessed 15 May 2010).
- Lonsdale, Jeremy and John Mayne (2005), 'Neat and tidy . . . and 100% correct: ensuring the quality of supreme SAI performance audit work', in Rob Schwartz and John Mayne (eds.), *Quality Matters: Seeking Confidence in Evaluation, Auditing and Performance Reporting*, New Brunswick, NJ: Transaction Publishers, pp. 173–98.
- Mayne, John and Rob Schwartz (2005), 'Assuring the quality of evaluative information', in Rob Schwartz and John Mayne (eds.), *Quality Matters: Seeking Confidence in Evaluation, Auditing and Performance Reporting*, New Brunswick, NJ: Transaction Publishers, pp. 1–20.
- Netherlands Court of Audit (2007), *Peer Review of the Netherlands Court of Audit. Report by the Supreme SAIs of Norway, New Zealand, South Africa and the United Kingdom*, The Hague: Netherlands Court of Audit.
- New South Wales Audit Office (2010), 'Guide to the Audit Office – summary', available at: http://www.audit.nsw.gov.au/info/hr/career_opportunities/summary.pdf (accessed 15 May 2010).

- Office of the Auditor General Western Australia (2009), *Annual Report 2008–2009*, available at: www.oag.wa.gov.au (accessed 15 May 2010).
- Office of the Comptroller and Auditor General (2008), *Peer Review of the Office of the Comptroller and Auditor General*, Dublin: Office of the Comptroller and Auditor General, available at: http://audgen.gov.ie/documents/miscreports/Peer_Review_2009.pdf (accessed 26 November 2009).
- UK NAO (2008), *Writing Smart Recommendations*, London: National Audit Office.
- UK NAO (2009a), *Value for Money Handbook: A Guide for Building Quality into VFM Examinations*, London: National Audit Office.
- UK NAO (2009b), *Annual Report 2009*, available at: <http://www.nao.org.uk/publications> (accessed 15 May 2010).
- US Government Accountability Office (GAO) (2007), *Government Auditing Standards*, Washington DC: Government Accountability Office, available at: www.gao.gov (accessed 4 January 2010).
- Victorian Public Accounts and Estimates Committee (2010), *Inquiry into Victoria's Audit Act 1994 Discussion Paper*, available at: <http://www.parliament.vic.gov.au/paec/reports/Inquiry%20into%20Victoria's%20Audit%20Act%20-%20V7%2020210%20to%20printer.pdf> (accessed 4 June 2010).

PART II

The contribution of performance audit

9. The impact of performance audits: a review of the existing evidence

Eddy Van Loocke and Vital Put

In this chapter,¹ the available literature about the impact of performance auditing is mapped and subjected to critical assessment. A conceptual framework is developed to reflect upon the impact of performance audits and 14 empirical studies are reviewed with the help of this conceptual framework. We then discuss the information on impact derived from the performance measurement systems of Supreme Audit Institutions (SAIs). Finally, we deal with the literature on the side-effects of performance auditing, and end with some practical conclusions and a research agenda.

INTRODUCTION

The task of the SAI has changed over recent decades. While in the past their main task was often to help their parliament to oversee government operations, as we saw in Chapter 1, an increasing number of SAIs now aim to contribute to a better functioning of government. This twofold mission is apparent from the websites of many SAIs. For example:

We hold government to account for the way it uses public money. We support, by helping public service managers improve performance. (National Audit Office, UK.)

The Algemene Rekenkamer aims at checking and improving the lawful, efficient, effective and ethical functioning of the state and its dependent bodies. (Algemene Rekenkamer, the Netherlands.)

It helps parliamentary assemblies and provincial councils to exercise control over the collection and use of government funds. . . . The Court seeks to improve the functioning of the state . . . (Belgian Court of Audit.)

Such assertions prompt the question as to the extent to which performance audits contribute to a more efficient and effective government. As an

increasing number of SAIs strive to establish a performance audit function, and this development is promoted in the developing countries by, amongst others, the World Bank, it is legitimate to consider the impact of this form of audit. In this chapter we therefore try to answer the following two questions:

- (1) What lessons can we learn from empirical studies on the impact of performance audit?
- (2) How do SAIs assess their own effectiveness?

The first question is divided into several parts. How do the studies define the dependent variables (impact variables)? What independent variables are supposed to contribute to the impact of performance auditing? What has been the methodological approach? In which conceptual frame can the research be placed? And to what extent are the findings of these studies transferable to other countries?

To answer these questions we have reviewed the existing empirical studies about the impact of performance auditing, a total of 14 in all (see Table 9.1). Studies about the impact of local SAIs² were not included in the analysis because there are important differences between local audit bodies and SAIs operating at a federal or regional level.³ The second research question was not dealt with as exhaustively, and we confined ourselves to compiling a range of typical impact indicators used by SAIs.

A FRAMEWORK TO REFLECT ON THE IMPACT OF PERFORMANCE AUDITING

Since performance audits have much in common with policy evaluations (see for example, Put, 2006), we relied on the literature on the impact of policy evaluations to develop a framework to review the literature on the impact of performance auditing. We drew a distinction between impact (the dependent variable) and the factors that facilitate this impact (the independent variables).

A Framework to Think About the Impact of Performance Audits

By impact we mean a wide range of direct or indirect, desired or undesired, influences that evaluations can exert. Weiss (1979) and Nutley & Webb (2000) distinguish the following forms of impact:

Table 9.1 Summary of the 14 reviewed studies

	Study design (country, period, methodology, variables, theory)
Alon (2007)	<p>Israel</p> <p>In 22 cases, dating back to the 1990–2002 period, among which 10 were from the SAI and 12 from journalists, 90 problems were pinpointed. Impact was measured by checking <i>to what extent the problems exposed had been solved and within what time frame</i>.</p> <p>Not theory-based</p>
Blume & Voigt (2007)	<p>Comparison between countries and their respective SAIs.</p> <p>The data are derived from 2 INTOSAI surveys (1995, 60 countries) and the OECD/World Bank (2003, 44 countries). Impact was defined in terms of <i>fiscal policy</i> (e.g. the size of the state budget deficit), <i>government effectiveness</i> (e.g. corruption perception index of Transparency International) and the <i>work – and capital productivity</i>. The independent variables were the SAIs’ mandate (type of audit), their independence (appointment) and the institutional environment (competition with private auditors).</p> <p>Economic framework</p>
Brown (1980)	<p>United States (federal government level)</p> <p>Survey at 7 SAIs: comparison of most and least effective audits (265 reports dating back from 1971–76).</p> <p>Effectiveness is defined as the <i>implementation of recommendations</i>. The author points at a number of influencing factors such as collaboration with the auditors, press interest, materiality of the findings, timing of the report, selection of the subject.</p> <p>Not theory-based</p>
De Vries (2000)	<p>The Netherlands</p> <p>3 reports (1985, 1989, 1993). Reconstruction of events with interviews and content analysis.</p> <p>Impact is seen as the implementation of <i>change</i>. This impact is, amongst others, dependent on the degree of cognitive coupling between auditors and auditees’ repertoires (shared ideas) and the degree of social coupling (distance between an auditor and an auditee).</p> <p>Social-psychological / constructivist frame of thought.</p>
Etverk (2002)	<p>Estonia</p> <p>3 reports dating back from 2001</p> <p>Questionnaires and interviews of auditors and auditees</p>

Table 9.1 (continued)

	Study design (country, period, methodology, variables, theory)
Etverk (2002)	Impact is defined in terms of <i>auditee's perception of the auditor</i> (satisfaction, collaboration, auditors' credibility), <i>auditee's perception of the impact on the audited entity</i> (added value, acceptance of recommendations) and in terms of <i>contribution to the public debate</i> (within the entity, in parliament, in the press). These impact indicators are related to a range of supposed determinants: the audit process (auditee's perception of: the auditor's willingness to collaborate, the power relations, the tone of the report, the auditor's competence), the auditee (level of commitment, fluid communication with the auditor, previous experience with the auditors) and environment factors (willingness at staff level, political will, timing, ongoing reorganizations). Social-psychological frame of thought, based on Morin (2001)
Hatherly & Parker (1988)	Australia (federal and regional) 6 reports dating back from 1984–86: 2 from the National Audit Office, 2 from the Victorian Auditor-General, 2 from the South Australian Auditor-General. The performance audit type (efficiency, effectiveness) was correlated with the auditee's reaction (<i>acceptance or rejection of the recommendations</i>). Not theory-based
Ingraham & Moynihan (2001)	United States (federal government level) Survey conducted in 50 states in 2000, content analysis of documents and interviews with officials. The survey is about the use of performance information, one aspect of which was if audit contributed to its use. Not theory-based
Johnston (1988)	United States (GAO) Analysis of 176, randomly selected recommendations (out of 1,134 recommendations in 1983). This survey is about the relation between the <i>acceptance of recommendations</i> and the type of recommendations. The type of recommendations is divided into four categories (processes, rules, structures, values), with increasing weight. Not theory-based
Lapsley & Pong (2000)	United Kingdom Interviews made in 1997–2000, of 12 experts in performance audit, in the private and public sector, locally and NAO. Not theory-based

Table 9.1 (continued)

	Study design (country, period, methodology, variables, theory)
Milgrom & Schwartz (2008)	<p>Israel</p> <p>17 reports dating from the period 1980–2000 on water policy were checked; the aim was to verify whether the problems reported in these audits had been corrected.</p>
Morin (2001)	<p>Political science framework</p> <p>Canada (OAG + Quebec)</p> <p>6 reports, from 1995 and 1996, 3 from the OAG and 3 from Quebec</p> <p>81 interviews of auditees (41), of auditors (38) as well as members of parliament (2).</p> <p>Qualitative content analysis of the interviews and documentation.</p> <p>Impact is operationalized in 14 indicators, in terms of <i>auditor's perceptions by the auditee</i> (attitude towards auditors), the <i>perception by the auditee of the impact on the organization</i> (added value, acceptance of recommendations, changes in management practices), <i>contribution to the public debate</i> (in parliament and the press). The independent variables consist of 9 indicators on the audit process (perception of: auditors' participative style, power relations, auditors' credibility, tolerance towards criticism, communication smoothness, etc.) and 7 indicators related to the environment (readiness of organization staff to cooperate, political will, audit timing, ongoing reforms in the audited entity or at government level, etc.).</p>
Morin (2004) / (2008)	<p>Socio-psychological framework</p> <p>Canada (Quebec)</p> <p>Survey with 99 (top and medium) managers and professionals from administrations audited by the Quebec provincial audit office. The survey covered audits implemented between 1995–2002; the number of reports is not known.</p> <p>The impact is operationalized in 10 indicators (audit added value, relevance of the recommendations, action taken following the audit, personal effects of the audit, etc.), as perceived by the auditee. In contrast to Morin (2001) no independent variables regarding the interaction between the auditor and audited were included. All factors are external: environment factors (readiness to cooperate of the base level, political will, timing, reforms), members of parliament's role, media role.</p> <p>Not theory-based</p>

Table 9.1 (continued)

	Study design (country, period, methodology, variables, theory)
Schelker (2007)	<p>United States (federal government level) Data collected between 1989 and 2006</p> <p>The study is about the relation between, on the one hand, SAI's institutional features (appointment /election) and the size of the performance audit function and, on the other hand, a number of macro-economic indicators (credibility, government performance index).</p> <p>The impact is measured by 3 indicators: <i>Moody's state credit rating</i> (a state market evaluation), the <i>government performance ratings</i> issued by the Governance Performance Project (these criteria refer to good management) and '<i>divided government</i>' (the extent to which the voter splits the power among the various parties (an auditor with a 'strong' mandate is assumed to have a reassuring effect for the citizen so that it is then unnecessary to split the power to weaken it). The independent variables were, on the one hand, the auditor independence (operationalized in the way he is appointed: election by citizens, appointment by parliament) and the scope of the performance audits (no performance audit, economy & efficiency audit, program audit, compliance audit).</p>
Van Der Meer (1999)	<p>Economic approach The Netherlands 2 reports (1987 and 1988)</p> <p>The author sees impact as the genesis of a <i>learning effect</i>. Socio-psychological / constructivist framework</p>

- Instrumental use: this assumes a linear process where research leads to knowledge and knowledge to policy.
- Conceptual impact (or enlightenment model): knowledge gradually permeates policy through all sorts of channels (for example, insights change slowly, theories are adapted and so on).
- Interactive impact: researchers constitute only one group of participants in the decision-making process among many others. The use of research is part of a complex process that also involves other information sources and existing insights and bias.
- Political-legitimizing use: knowledge is an asset in a political debate (for example, the opposition in parliament can use an evaluation report to establish that a policy is not effective).

- Tactical use: knowledge is used to influence a decision-making process (for instance, a policy decision can be put off awaiting further evaluations).

Not only can the evaluation report generate an impact, but the evaluation process itself can also exert an influence (Patton, 1997; Kirkhart, 2000). And impact is not merely instrumental; the communication between an evaluator and the stakeholders during the process can bring about clearer insights into the issue and changes in the relations between persons and entities. An impact can thus take various forms. It can also take place during the evaluation, or shortly after the communication of the evaluation results, but it can also be felt a long time later. The impact is not always measurable (for instance, the conceptual impact of an audit) because, for example, of the long elapsed time and the complicated causal pathway between the evaluation and the impact. Finally, an evaluation can have an impact that is desired or not desired by the evaluator.

Based on these insights a framework was worked out to map the impact of performance audits (see Table 9.2). Column one categorizes the types of impact, on the basis of Weiss' model. In the second column these types are further subdivided into specific forms of impact arising from performance audits. The third column indicates the time dimension within which the impact takes place. The fourth column contains an overall assessment of the measurability (including the possibility of establishing a causal link) of the impact, based on our professional experience and on the reviewed literature. Some types of impact are hardly measurable because the impact takes place in the long run and the causal relationship between a change and the audit is hard to prove. The time dimension and measurability can influence the selection of indicators to measure the impact and thus indirectly the way we look at the impact.

The different types of impact can influence each other. For example, it can be assumed that members of parliament exercise pressure on the minister to implement certain recommendations, interest groups can exert an influence on members of parliament and cabinet, media interest can raise the interest of study services for the report, and so on. These forms of impact can then be considered as intermediary variables starting off other forms of impact.

This framework based on evaluation literature has two limitations. Firstly, literature about the impact of evaluations mainly focuses on the impact on improving government effectiveness and efficiency. As we have seen, however, SAIs have a twofold task. The accountability aspect receives less attention in the evaluation literature (although see Bemelmans-Videc, Lonsdale and Perrin, 2007). Secondly, the framework

Table 9.2 Types of impact of performance audits

Type of impact	Effect of a performance audit	Time dimension	Measurability
Instrumental	<p><i>Anticipation on the outcomes</i> Action is often taken during the audit process to remedy failings found in the functioning of a government entity. Often it refers to an adjustment of processes. The auditee anticipates the audit findings.</p>	During the audit process	Good
	<i>Acceptance of recommendations</i>	Short term	Good
	<i>Implementation of recommendations</i>	Short, medium to long term	Reasonable
	<p><i>Changes not directly subsequent to the recommendations and side effects</i> These changes can be widely varied. The auditee can solve the malfunction in a way other than that presented in the report, and the report can be used or misused in the evaluation process of public managers. Moreover, the results may be used by actors who were not subject of the audit. This makes sense because reports sometimes contain a generalizable standards frame of best practices.</p>	Short to long term	Difficult
Conceptual	<p><i>Learning process and change in mental or intellectual frames</i> Conceptual impact can emerge both during the process itself and after the publication of the report.</p>	From process to long term	Difficult

Table 9.2 (continued)

Type of impact	Effect of a performance audit	Time dimension	Measurability
Interactive	<i>Consultation and negotiation between stakeholders</i> This impact can relate to consultation between the administration and actors from the civil society, such as a target group, and the office of the minister or between the office of the minister and interest or target groups.	From process to medium term	Difficult
	<i>Impact on other 'knowledge generators'</i> Audit reports can supplement other information held by think tanks, universities, study services and so forth.	Short to long term	Difficult
Political–legitimizing	<i>Media interest</i> The impact from the media can be accommodated in the legitimizing impact for two reasons. First, the interest from parliament is likely to be raised by interest from the media. Second, media are often used to legitimate or criticize a policy after completion of an audit.	Short term	Good
	<i>Interest from parliament</i>	Short term	Good
Tactical	<i>Impact on a changing or decision-making process</i> Tactical use of an audit can take various forms, e.g. an audit envisaged by an SAI can be a prompt for the minister to postpone a decision.	During the process and short term	Reasonable

is applicable to individual audits. However, the mere existence of an SAI and its mandate could by itself have an impact. This receives no attention in the evaluation literature (but some impact studies on auditing try to capture this, as we will see).

A Framework to Think About the Factors that Facilitate the Impacts of Audits

Table 9.3 contains a model showing factors likely to contribute to the impact of performance audits. The model is based on recent literature on the use of evaluations and research.⁴ The model is structured around three levels. The *micro level* involves factors specific to an individual audit and we distinguish five clusters of factors. The literature on research utilization focuses strongly on these factors to explain the use of scientific knowledge.

The *meso level* relates to the SAI's characteristics and the audited entities' characteristics. The trust an evaluator enjoys from the evaluated entity, which is closely linked to the evaluator's reputation, is an important factor. Power and institutional anchorage are among the factors specific to SAIs. The literature stresses the importance of the features of the knowledge provider, but does not refer explicitly to powers. We think this can constitute an important factor for the usability of audits, specific to audit institutions. The way priorities for evaluations are determined and the resources available for evaluation have also a key influence on the extent to which a policy is supported by scientific knowledge (Campbell et al., 2007). These factors are not mentioned here because, as we saw in Chapter 3, audit institutions select their audits themselves. Transposed into performance auditing it relates to the quality of the selection procedure in audit institutions. More specifically, we are convinced that the selection procedure should result in the implementation of policy relevant audits. It makes a difference for an auditee if evaluation and drawing lessons are considered as a logical link in the policy and management cycle and both the required management instruments and set-ups have been developed (for example, performance measurement systems, specialized units, reporting procedures, and so on). This is closely related to the cultural component: to what extent is the auditee open for evaluation and audit? Does the auditee support the principles of monitoring, evaluation and auditing? Have officials and policy makers internalized these values?

The *macro level* includes some characteristics of the public sector. We did not find much literature about the influence of the macro level on the utilization of evaluations or scientific knowledge. From Varone and others (2005) it appears that the way the political system is organized

Table 9.3 Factors facilitating impact

Level	Type of factors	Factors that facilitate impact
Micro level (1)	Report	<ul style="list-style-type: none"> ● Accessible to the actors who are the addressees
	Process	<ul style="list-style-type: none"> ● Summary with recommendations ● Audit planning takes the societal and administrative context in which the audit will be performed into account ● Good communication during the process: from preparation – insight into the mental models of audited entities – to the reporting ● Personal contact between auditors and auditees ● Good methodology ● Objective ● Reliable sources of findings ● Quality control (for example, peer review)
	Diffusion and communication	<ul style="list-style-type: none"> ● Use of a range of means of communication such as press releases, presentations, workshops, briefings, and face-to-face contacts ● Diffusion and communication of results not only to the auditee, but also to all other actors concerned ● Addressee’s tailored communication
	Audit results (findings, conclusions, recommendations)	<ul style="list-style-type: none"> ● Results are concomitant with other knowledge – the new set of knowledge is ‘acceptable’ for the auditee ● Increase in quantity of unambiguous knowledge has a booster effect ● Decrease in uncertainties with the auditee and testing of the assumptions with the auditee on the topic (Leviton’s mental models) ● Results fit in the overall context ● Timely to generate or influence a decision ● Relevant for the policy and the management ● Knowledge fits in the frame of the auditee’s present concern and activities
	Reaction from third actors	<ul style="list-style-type: none"> ● Clear, usable and feasible recommendations ● Presence of advocacy coalitions (among others role of parliament as SAI’s ‘customer’)

Table 9.3 (continued)

Level	Type of factors	Factors that facilitate impact
Meso level (2)	SAI	<ul style="list-style-type: none"> ● Enjoys auditee's trust ● SAI's reputation ● Selects relevant topics ● Publication policy – development of diffusion channels and communication methods ● SAI's powers (mandate) and institutional anchorage
	Auditee	<ul style="list-style-type: none"> ● The use of scientific knowledge (for example, procedures, knowledge management) is structurally embedded ● Nature of the policy domain (for example, familiarity with evaluation and audit, technical character of the policy domain, autonomy of the auditee) ● Attitude towards knowledge, audit-minded or not audit-minded
Macro level (3)	The policy – institutional system	<ul style="list-style-type: none"> ● Institutional anchorage of evaluation (for example, obligation to assess policy on a regular basis) ● Characteristics of the political system (for example, position of parliament towards government, party discipline, decision-making processes)
	Values as far as management and policy are concerned	<ul style="list-style-type: none"> ● Attitude towards knowledge, evaluation culture

Notes:

- (1) Weiss and Bucuvalas 1980, Lindquist 1990, Oh 1997, Nutley et al. 2002, Hanney et al. 2003, Sutcliffe and Court 2005, Innvaer et al. 2002, Campbell et al. 2007, Leviton 2003, Johnson et al. 2009.
- (2) Sutcliffe and Court 2005, Campbell et al. 2007, Brans and Van der Straeten 2004, Oh 1997, Rossi et al. 1999, Bongers 2009.
- (3) Oh 1997, Brans 2004, Bongers 2009, Varone et al. 2005.

(administered centrally or federated, or the power relation between the executive and the legislature) plays a major role in the significance of evaluations and scientific knowledge for policy decisions.

The literature is not unambiguous about the importance of the factors. Sometimes they are aggregated (for instance, report 'quality'), or arranged

differently from the model used here. However, the factors identified show a large degree of consistency. Timeliness and relevance of results are strongly dependent on the organization of the audit process. Advocacy coalitions are only possible if stakeholders know the results of the audit and if sufficient coverage has been given, for instance via the media or through direct contacts between auditors and stakeholders. Moreover, several factors which influence the usability in terms of individual audit are determined by the audit institutions' characteristics and the characteristics of the whole public sector. It is taken for granted that a strong parliament is more likely to boost the impact of audits with the administration and with the minister than a parliament that is hardly willing or able to control the executive.

WHAT DO WE KNOW ABOUT THE IMPACT OF PERFORMANCE AUDITING?

A Review of the Empirical Research on the Impact of Performance Auditing

To find studies on the impact of performance auditing we screened the literature (journals, books) as well as the grey literature (congress papers, theses). We found 14 empirical studies (see Table 9.1).

How do impact studies define the notion of 'impact'?

Table 9.4 uses the impact categories from Table 9.2 and summarizes how many among the studied publications discuss these categories. The number included in this table is 10 because in three publications the impact is not examined in terms of audits, but in terms of the organization and the SAIs' overall functioning. In one publication the notion of 'impact' is not operationalized.

In nine out of the 10 publications in Table 9.4, impact is described as instrumental. Other types of impact are discussed in a limited number of these publications or not at all. This is in line with the greater attention usually paid to the instrumental impact of evaluations. The SAI's focus on improving the functioning of the public sector probably also contributes to an instrumental impact approach. A performance audit aims at verifying an existing situation in the light of an alternative and better one. The gap between the real situation and the desired situation needs to be filled and SAIs frame recommendations to this end. Based on this audit model, the ideal impact is therefore instrumental and, more particularly, the implementation of the recommendations. Most authors frame impact

Table 9.4 What types of impact do impact studies pay attention to?

Impact type	Effect of a performance audit	Number of empirical studies discussing this dimension (N = 10)
Instrumental (9 publications)	Anticipation on outcomes	None
	Acceptance of recommendations	3 publications
	Implementation of recommendations	1 publication
	Changes achieved	6 publications – operationalized as follows: <ul style="list-style-type: none"> ● Remedy malfunctions ● Time interval between finding and setting right malfunction ● Change resulting from audit ● Changes brought about by the auditee with organizational or personal effects ● Influence on management practices ● Influence on relations between the auditee and interest groups
Conceptual (3 publications)	Learning process and change in the mental or intellectual frames	3 publications – operationalized as follows: <ul style="list-style-type: none"> ● Change in frames, repertoires ● Debate in audited entity ● Learning effects
Interactive (1 publication)	Consultation and negotiation between stakeholders	1 publication
	Impact on other ‘knowledge producers’	1 publication
Political legitimizing (3 publications)	Media interest	2 publications
	Interest from parliament	2 publications
Tactical (0 publications)	Impact on a changing or decision-making process	None

in terms of the achievement of changes, whether or not derived from the recommendations. De Vries (2000) and Van Der Meer (1999) indicate that defining usability in terms of implementation of recommendations is not adequate because a causal connection between a report and a subsequent change is difficult to establish. None of the studies paid attention to anticipative use of auditing findings during the audit, which is also a form of instrumental use. Nevertheless, from experience we know that the auditee will often take action before the report has been drawn up. It is in an auditee's interest to ensure that malfunctions have been corrected before the final report is publicized so that deficiencies are either not reported or publicized at all, or done so less conspicuously. Bongers (2008) hints at a behaviour that would be similar to the well-known Hawthorne-effect in occupational psychology: merely drawing attention to a process or activity has the effect of ensuring a smoother completion of the process or activity.

Only a minority of the publications explicitly consider the other forms of impact. These forms of impact are often considered as a determinant of an instrumental impact. Morin (2008) stresses that audit outcomes permeate slowly and subtly, and she views conceptual use as an intermediate stage leading to instrumental use. Etverk (2002) is the only one to hint at an interactive use in stimulating the debate with third parties. De Vries (2000) and Van Der Meer (1999) state that interactive use is an essential condition for impact.

In five publications, interest from the media and parliament is seen as having a big influence in achieving an (instrumental) impact. A significant task for SAIs is also to contribute to parliament's oversight function. But only two publications examined the political use of audits explicitly as a form of impact.

The tactical use of audit results – such as commissioning a new evaluation, or the postponement or acceleration of decision-making – was not found in any publication as a form of impact. There is, however, anecdotal evidence that audits are used tactically. Radcliffe (1997) illustrated how an audit of the provincial audit office of Alberta in Canada, implemented at the government's request, had been tactically used by politicians to refute criticism on their policy and to blame management. It was noted (pp. 358–59):

In having auditors investigate these matters, politicians brought to bear an approach and analysis which would encourage an understanding of the events surrounding NovAtel in managerial rather than political terms. . . . The auditors' diagnosis, as read in their report, was entirely managerial: it approached causality in terms of formal rationality alone . . . The Minister for Technology, who had earlier promised to resign in a minute if found guilty of wrongdoing,

carefully noted that 'He [The Auditor General] has not . . . fingered me in any specific way in the report'. However, in claiming that NovAtel's management had been deceiving everybody (. . .) cabinet ministers depicted themselves as innocent bystanders, misled by a corrupt and nameless management.

Morin (2001, 2008) and Etverk (2002) consider the auditee's perception and reactions related to the auditors as impact (dependent variable). This, however, seems an impact determining factor, rather than an impact itself.

Schelker (2007) and Blume and Voigt (2007) developed a new perspective on the impact of performance audits and examined the impact on a macro level: including the relationship between the features of SAIs and macro-economic indicators (such as size of the state budget deficit, or credit ratings), corruption levels and governance indicators, and so on. These studies refer more to impact related to the SAI characteristics than to the outcomes of an audit process. Similarly, reference is often made to a preventive impact resulting from the mere existence of audit offices and the possibility that an audit could be implemented (Morin, 2008; Brown, 1980; Ingraham and Moynihan, 2001).

This review shows very clearly that impacts are typically described as instrumental but that authors also identify a range of forms of impact. Furthermore, the difference between impact and impact-determining factors is vague in the 14 reviewed studies. Nearly all publications take into account the long time lag between the publication of a report and the changes introduced within the audited entities. In the case studies, selected audits nearly always date from four years or more before. The reviewed studies focus on the impact of performance auditing in the sense of contributing to a more efficient and effective government. It is somewhat overlooked that performance audits can also contribute to improving accountability. There is no empirical examination of to what extent and under what circumstances the efforts to ensure a better functioning public sector and assisting parliament in its oversight function are compatible objectives.

Do audits have an impact?

Figure 9.1 shows the availability of information on the actual impact.

Three publications try to quantify the impact of audits (see Table 9.5). Schelker (2007) also quantified the impact, but at a meso level: he found a positive and significant link between, on the one hand, the existence of a performance audit function and the scope of these audits and, on the other hand, the scores on the 'government performance rating' (a scale that measures the existing good management practices, for instance in the field of HRM management, IT management).

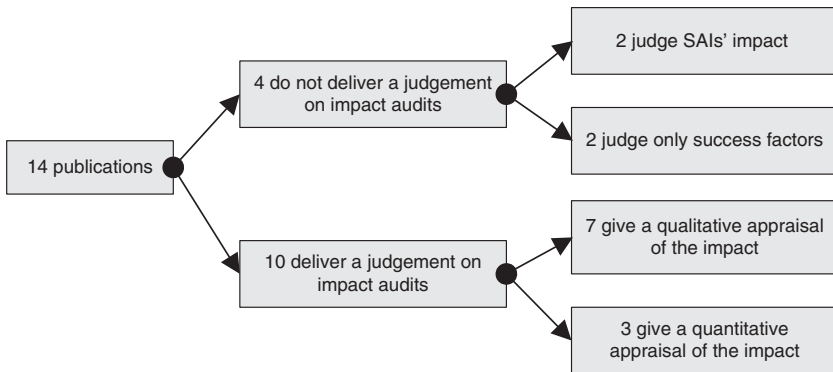


Figure 9.1 Impact information in the 14 reviewed studies

The other studies provided a qualitative appraisal of a limited number of cases:

- Morin (2001) found that out of the six cases examined, two were successful, two had little success and two were fairly successful.
- Brown (1980) compared seven successful with seven unsuccessful audits in the United States (the selection criteria were not clarified).
- On the basis of a range of interviews with VFM specialists in the United Kingdom, Lapsley and Pong (2000) found that performance audits have an operational and a strategic impact according to their respondents. The exact meaning of ‘impact’ was not clarified in the publication.

Table 9.5 Quantitative information on impact (at audit level)

Publication	Results
Alon, 2007 – Israel (10 reports between 1990 and 2002)	46% of malfunctions are corrected, or, if we include those almost corrected, 57%. (Time lag between finding of the malfunction and correction is about 3.8 years)
Milgrom & Schwartz, 2008 – Israel (17 reports between 1980 and 2000)	49 of 278 malfunctions (17.6%) were corrected as a direct consequence of the reports. 134 of the 278 malfunctions (48%) were corrected, although the link to the report was not evident
Johnston, 1988 – USA (176 recommendations from fiscal year 1983)	Between 51% and 77% recommendations were accepted

In a nutshell we can conclude that the publications analysed, the qualitative studies as well as the quantitative, show that performance audits have an impact. The picture, however, is varied: some audits appear to have significant impact, while others have none or only a partial impact.

Various studies support the view that the impact of performance audits is often slow to materialize and does so in a complicated way. Morin (2008) suggests that audits do not bring about radical changes in the audited organizations, but have a slow yet perceptible influence. De Vries (2000) and Van Der Meer (1999) consider that policy changes or changes in organizational practices hardly if ever have only a single cause. They refer among others to conceptual and interactive forms of impact.

When measuring impact, underestimation is more likely than overestimation. Much information collected in case studies comes from the auditee and is corroborated with other information sources. It is possible that auditees underestimate the influence of audits, whether knowingly or not. As Mark Funkhouser argues, experienced auditors believe they know that audits can have an impact through the transfer of lessons from an agency on the subject of performance audit to other agencies that face similar challenges (for example, how to deal with major investments). This conceptual form of impact is hardly measurable in a scientific study. It supports the idea of underestimation of the global impact of performance audits.

What were the facilitating factors examined in the impact studies?

Table 9.6 takes up the facilitating factors from Table 9.3 and identifies the number of publications in which they are cited.

This table shows several important blind spots in the existing studies on the impact of performance auditing. These are: the quality features of the report (such as the readability and language use, and availability of a summary); the technical aspects of the audit process (such as the methodology used); the diffusion of the report or results (press release, convocation activities) and the political-institutional system and values and culture (except for the concurrence of audits with reform processes within government).

What facilitating factors contribute effectively to the impact of audits?

What are the factors which have a proven and significant impact on audit use? The research above concludes that all factors listed in the previous paragraph play a role to some extent. A more refined analysis shows that a few factors are more recurrent and certain authors attach much weight to them:

Table 9.6 What facilitating factors receive attention in the impact studies?

Level	Cluster factors	Number of empirical studies where this factor is discussed (N = 11)
Micro Factors specific to an audit	The report	None
	The process	4 publications – they stress most of all a good relationship between auditors and auditees. Auditing is a process of social interaction. Technical aspects related to the quality of the examination process (for example, methodology, reliable sources, quality control) are not looked into
	Diffusion and communication	None
	Results	8 publications – the factors examined are closely connected to the literature on policy evaluation and evidence-based policy (relevant and timely results, connected with existing knowledge and conceptions)
Meso Factors specific to the SAI and the audited entity	Third-party actors (including media)	5 publications – besides the great importance they attach to the influence of interest groups, they also emphasize the influence of media and parliament
	The SAI	3 publications examined whether an SAI's characteristics influenced the use of their audits. In two cases they refer to the SAI's organizational features or powers. The SAI's ideas on management and policy can have an influence. The SAI's reputation and policy were not examined
	The auditee	The auditee's characteristics were discussed in 5 publications. They did not deal so much with knowledge management but most of all the attitude towards an audit (including the experiences with previous audits) and the existence of reform initiatives within the auditee
Macro Factors specific to the politico-institutional context	The political –institutional system	3 publications paid attention to the role government reforms can have for encouraging the use of audits. The political system and institutional anchorage of evaluation were not discussed
	Values and culture	None

1. The *relationship between auditors and auditees*. To work well, this relationship has to be based on a good communication between both parties, in a context of mutual trust. Both parties have to behave like partners and the audit has to be implemented in a constructive context. De Vries (2000) stresses the importance of ‘social coupling’ between an auditor and the auditee. Moreover, a previous experience with audits in the audited organization can have either a positive or a negative influence (Etverk, 2002).
2. The characteristics of the audit *results* (findings, conclusions, recommendations) are cited on numerous occasions as factors influencing its use. The audit has to be relevant for the auditee, and the results have to be connected with the auditee’s priorities and ongoing reform initiatives. The timeliness of the results is linked to this. The audit is likely to be more influential if the conclusions and recommendations are not too radical (Johnston, 1988) or if there is a cognitive coupling of repertoires, as it might be expressed in social constructivist terms (De Vries, 2000). At its most simple, this means that auditors and auditees have to share ideas. Hatherly found that recommendations aimed at increasing effectiveness are followed less than recommendations aimed at a more efficient operation (Hatherly and Parker, 1988).
3. As to the role of *third parties*, authors emphasize the positive influence of parliament, as well as media interest in the audit results. Interest from parliament has both an intermediary or indirect impact, as well as a direct impact. An audit institution contributes directly to parliament’s oversight role and the impact is intermediary in the sense that pressure from members of parliament can result in the implementation of recommendations (instrumental impact). According to Morin (2004) the media’s influence can also have side-effects such as the discouragement of officials. Only a few authors also highlight the influence of interest groups in ‘advocacy coalitions’.
4. The *evaluation culture* in the audited entity seems to play an important role, and in particular, whether or not the audited entity is open to advice.
5. The coincidence of *reform projects* within the auditee and the audit results can also have a favourable effect. Such a concurrence creates a window of opportunity for the audit.

Differences in *SAIs’ institutional and organizational structures* do not show a significant relationship, although Schelker (2007) is of the opinion that some evidence shows that, within the US context, it is preferable to elect the head of SAIs instead of appointing them. Chance events can have a positive or negative influence on the impact.

Several studies mention chance events as being an explanatory factor for the audit impact (Milgrom and Schwartz, 2008; Morin, 2004). For example, a report on climate policy is probably likely to arouse more interest in the media and parliament just after a heat wave or a flood than during mild winter months.

Is it possible to generalize from the impact studies examined?

The findings in the 14 impact studies are not transferable as such across SAIs because of a range of methodological and contextual constraints. Thus, apart from the studies by Johnston (1988) and Hatherly and Parker (1988), which look explicitly into the relationship between the nature of an audit and the extent to which recommendations have been implemented, there is otherwise no explicit reference to what kind(s) of performance audits are involved in the analysis of the impact. Nevertheless, performance audits are very heterogeneous products, even within one SAI. It is plausible that there is a link between the nature of the implemented performance audits and the degree and nature of the impact (for instance, an audit that focuses on processes is probably more likely to result in a larger and quicker utilization than an audit focused on – less easily changeable – structures). It also makes sense to anticipate that utilization is different in the case of an audit implemented at the request, for example, of parliament or the government, rather than one identified on the basis of an SAI's own planning. This variable is not studied in any study. Since nearly all studies tackle all performance audits as a whole, these links are not made visible.

A second factor is that 8 of the 14 studies are based on case studies, sometimes merely cases from one specific policy domain. This also restricts the potential for generalization. Thirdly, the small number of studies, some of which are more than 20 years old,⁵ also reduces the possibility to formulate general conclusions. Fourthly, the institutional and social contexts (including the powers of an SAI, parliamentary system, status of the media, and role of civil society) vary from country to country but the set of studies concerned only a limited set of countries.⁶ Fifthly, most studies use vaguely defined variables. Finally, the possibility of aggregating the research findings is limited by the fact that studies proceed from differing conceptual frameworks: seven are not theory-oriented (they proceed mostly from impact indicators such as the rate of implemented or accepted recommendations or the rate of problems solved), four proceed from a social-psychological framework (with a strong focus on the audit process), two proceed from an economic framework (with a strong focus on macro-economic variables) and one starts from a political science framework. Within these frameworks

other variables are often used. The study results are thus not easily comparable.

Information from SAIs' Performance Measurement Data

SAIs increasingly measure their own activities.⁷ There is also an increasing interest in performance measurement within INTOSAI, as evidenced by the fact that it recently organized a symposium on this topic⁸ and subsequently a working group was set up.⁹ The SAIs' performance indicators cover all aspects of what has been termed the MAPE spectrum illustrated in the following examples:

- **Means:** cost of an audit, work time spent to complete an audit, amount of resources allocated to *front line* activities.
- **Activities:** implementation of audits within a planned timeframe, number of performance audits for which external professionals are called upon.
- **Performances delivered:** number of audit reports produced, timely delivery of audits.
- **Effects:** number of implemented recommendations, generation of a debate in parliament and media.

As far as this chapter is concerned we are mostly interested in the effect indicators. From a survey organized by the European Court of Auditors¹⁰ it appears that European SAIs use the following effect indicators (the figures between brackets refer to the number of SAIs that are using this indicator, n = 24):

- Implementation of audit recommendations (12)
- Auditees' views on the value of the SAI's work (8)
- Number of SAI's audit reports discussed in parliament (6)
- Parliament's views on the value of the SAI's work (5)
- The number of times an SAI is featured in the media (4)
- The financial impacts of an SAI's audits (3)

To take this further, Table 9.7 illustrates how a few European or other SAIs use different types of impact indicators.

The percentage of 'implemented / accepted recommendations' is, according to the European Court of Auditors' survey, the most commonly used indicator. A somewhat older survey conducted by the World Bank/OECD (World Bank, 2003) also revealed that many SAIs were using this indicator, and it is therefore relevant to draw attention to important validity constraints. These include:

Table 9.7 Types of impact indicators used by SAIs

Type indicator	SAI	Indicator	Score
Implementation of audit recommendations	Office of the Auditor-General of Canada ¹	Percentage of performance audit recommendations reviewed that are endorsed by the Public Accounts Committee. Percentage of performance audit recommendations implemented four years after their publication	In the year 2006–07 the rates for both indicators were respectively 74% and 44%
Auditees' views on the value of the SAI's work	Office of the Auditor-General of Canada ²	Percentage of departmental senior managers who found our performance audits add value	60%
Number of SAI's audit reports discussed in parliament	Belgian Court of Audit	Number of performance audit reports discussed in the Flemish Parliament	In both 2006 and 2007 91% of reports and in 2008 100% of reports were discussed in the Flemish Parliament
Parliament's view on the value of the SAI's work	Australian National Audit Office ³	Parliament acknowledged the value of the ANAO contribution	93% of parliament members surveyed expressed satisfaction with the ANAO products
Number of times an SAI is featured in the media	Belgian SAI	Media interest for audit reports	Qualitative indicator – evaluated by audit and not aggregated

Table 9.7 (continued)

Type indicator	SAI	Indicator	Score
The financial impacts of an SAI's audits	National Audit Office ⁴	Financial impacts achieved ⁵	In budgetary year 2007–08 a financial impact amounting to £9.28 was achieved for every £1 of NAO's operating cost

Notes:

¹ Office of the Auditor-General of Canada, 2006–07 Estimates Report on Plans and Priorities.

² Idem note 12.

³ Australian National Audit Office, Annual Report 2007–2008.

⁴ National Audit Office, Annual Report 2008; Haynes Watts (2004).

⁵ This comprises savings on expenditure, efficiency improvements and improvements in revenue collection. The NAO has developed a methodology to map its financial impact. These figures are validated by the audited department, verified by the NAO internal auditor and also subject to an external audit.

- The impact of an audit goes further than the mere implementation of recommendations; there is also an indirect use through influencing ideas (conceptual impact).
- Some improvements may have already been carried out during an audit, in which case there is no need to formulate any recommendations. This anticipatory action is not reflected in the figures, although the impact is certainly there.
- The implementation of a recommendation does not necessarily need to result in an improvement (some recommendations may not be effective), and not implementing a recommendation is not necessarily bad (a problem can be solved in another way).
- The percentage of recommendations implemented is strongly correlated with the nature of the proposed recommendations. In the case where an SAI suggests the mostly easily implementable changes, the percentage of implemented recommendations will be high. The more radical a recommendation, the lower the likelihood of implementation, a point shown by several impact studies such as those by Hatherly and Parker (1988) and Johnston (1988). Thus, strong emphasis on this indicator in the performance measurement system of an SAI could encourage the picking of 'low hanging fruit'.

- Definitions of this indicator vary from country to country; time-frames used to check whether recommendations have been implemented differ (2, 3, 4 years); some SAIs distinguish between fully and partially implemented recommendations; and some proceed from the percentage of recommendations that were agreed by the auditee, others focusing on those that were implemented. As these definitions are different and, as mentioned above, the nature of the recommendations differs strongly, it is hard to compare the percentage of implemented recommendations between countries.

The information obtained from SAI performance measurement systems is only partially comparable with the information derived from the impact studies on performance auditing discussed earlier. This is because other indicators are also used, such as those which relate to the SAI's accountability mission (for instance, the number of the SAI's audit reports discussed in parliament, parliament's views on the value of the SAI's work, and so on), which are less commonly found in the impact studies. Nor is an indicator such as audit's financial impact to be found in the impact studies. As a result, this type of information constitutes a meaningful complement to the impact studies.

However, there is one type of indicator that is found in the SAIs' performance measuring as well as in the scientific impact studies, namely the instrumental use of performance audits, measured by indicators such as the percentage of implemented recommendations. What is striking here is that scientific studies show a rather moderate impact (46 per cent completely solved problems or 57 per cent completely and nearly completely solved problems according to Alon 2007; 17.6 per cent according to Milgrom and Schwartz 2008), while SAIs' performance measurement systems often show a stronger effect.¹¹ How do we explain this difference?

No final answer can be given on this in view of the limited information available, but we suggest two hypotheses.

- (1) The difference could be attributed to the fact that the figures relate to different work (different SAIs, different points in time, different reports, and so on) and consequently are not comparable.
- (2) Another explanation could be that SAIs themselves measure goal achievement (with no necessary causal relationship) while (a certain number of) studies (endeavour to) measure effectiveness, which implies a causal relationship between an audit and a solution to a problem. As a matter of fact, certain shortcomings found in audits can also be solved without an audit. Morin (2008, 697) as well as Milgrom and Schwartz (2008, 869), who both rely on the perceptions

of audited entities to prove a causality, indicate that the auditees can find it advantageous to claim that they would have solved the problems as well by themselves without an audit.

Literature on the Side-Effects of Performance Auditing

Finally, we turn to an issue that is covered in more detail by Frans Leeuw – the side-effects of audit. There is hardly any *empirical* research on the side-effects of SAIs' performance auditing. Only Morin (2008), based on a survey, mentions a few side-effects and also puts into perspective their importance. Other literature¹² is hypothetical by nature. Moreover, this literature lacks a shared and systematic approach, developed typologies or models, and the connection between these side-effects and their causes is barely discussed. The present literature does not go much further than making some sort of risk analysis of what could go wrong. The side-effects cited in this literature include:

- Auditees could create an *illusory world* made up of beautiful plans, SMART objectives, indicators, procedures, and so on, to keep the auditor satisfied, while behind this formal façade everything goes on as it did before. 'Decoupling', 'window dressing', 'dramaturgical performance', 'impression management', 'gaming' are all terms used to describe aspects of this phenomenon.
- *Narrowing of vision* ('tunnel vision', 'myopia', and various variants such as 'short-termism'). Auditors emphasize certain aspects of performance, while others are not examined and therefore risk being overlooked. Examples are the emphasis placed on efficiency and overlooking effectiveness, highlighting the rational-technical side of entities and neglecting the human side (culture, social capital, trust, incentives), or focusing on performance without checking the legality.
- *Suboptimization*. Examples are the improvement of a subsystem to the detriment of an organization as a whole, or the improvement of one organization in a policy field to the detriment of the policy field as a whole.
- Audits that emphasize formalization, procedures, and so on, can cause *rigidities*, the stifling of initiative and risk aversion ('ossification', 'manualization', 'analysis paralysis', and 'trust killers' are terms used in this respect).
- An audit requires the audited entities' collaboration and thus entails a *financial cost* for the auditee. One question is whether this cost outweighs the perceived benefit derived from the audit.

The literature puts forward a few reasons to explain the (possible) occurrence of side-effects. SAIs have a monopoly position, which may mean they are less sensitive to costs imposed on the auditee. Some features of the performance audit process are supposed to contribute to side-effects, that is, one-sidedness (for example, single-sided attention for systems and procedures; insufficient attention to the human side of organizations) and the style of the audit work (a negative, fault-finding approach). The striking fact is that these hypotheses focus exclusively on the auditor. Yet it seems likely that the features of the auditee and the institutional context could also contribute to the occurrence of undesired side-effects (for instance, a failing learning ability and criticism aversion within the auditee can also cause rigidities and excessive risk aversion). To date, however, we have little or no empirical evidence to help us understand whether or how these factors work.

CONCLUSIONS

This review of the empirical literature concludes with a positive message: performance auditing appears not to be a waste of time and has value. As Funkhouser elaborates further in the next chapter, based on evidence from auditors themselves, performance audits often have an impact. This may be ‘slow and subtle’ as one observer has concluded (Morin, 2008), or at times quick and explicit. At other times, however, performance audits seem to have had hardly any impact. In most reviews ‘impact’ is predominantly defined as instrumental, whilst a range of other impact dimensions (political in terms of legitimizing use, tactical use and so on) receive little attention. Non-instrumental impact types such as conceptual or interactive impacts are often considered as intermediate stages or success factors for instrumental impact.

What factors do contribute to an impact? Much uncertainty still exists about this. Some variables are hardly mentioned in the existing research (including features of the political-institutional system, the quality of the audit process, and the diffusion of the audit results). We do not know whether the impact is stronger when audits are carried out at the request of parliament or are initiated by SAIs, or about the causal relationships that may exist between a range of factors and the impact they may have (for instance the relationship between media and parliamentary attention and its impact). The nature and the weight of the causal relationships is not considered in this research.¹³ Nor do we know whether, or to what extent, SAIs’ dual role as watchdog and counsellor has an influence on the impact of audits. Nevertheless, the existing research tells us something about what

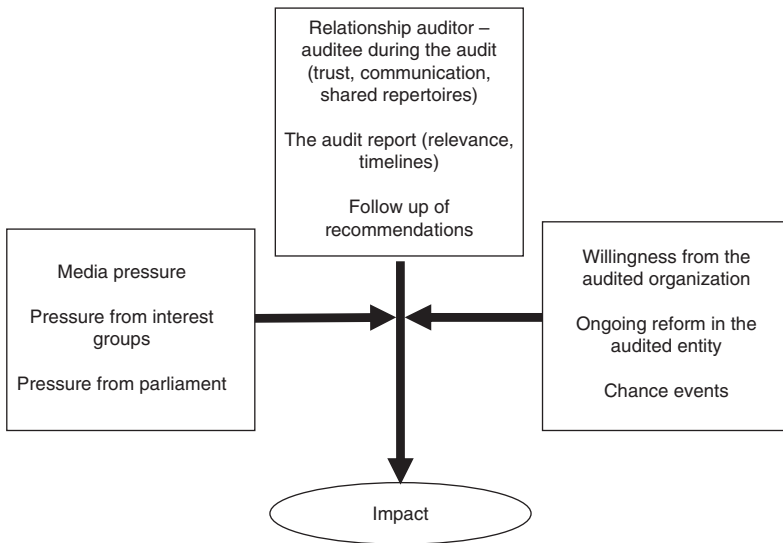


Figure 9.2 Factors that determine the impact of performance audits

factors contribute to an impact. Figure 9.2 represents the most meaningful causal relationships.

So how can SAIs use this knowledge to reinforce their impact? Based on our work, we consider:

- They can optimize factors that are within their control such as the auditor–auditee relationship. There can be a dilemma here, however, resulting from the SAIs’ double mission: the independent oversight function is more often characterized by social distance and mistrust; utilization of recommendations, however, requires social proximity and trust, a dilemma considered by Morin (2003) and Saint-Martin (2004).
- For factors that are outside their direct range of influence they can develop influencing strategies. For example, an SAI can manage its reputation, can try to find ‘advocacy coalitions’ with research institutes or the media.
- If SAIs are keen to make relevant recommendations while the audit topic is in the spotlight, they need to invest sufficient resources in the audit selection and planning. However, due to the time elapsed between the selection of topics and an audit completion, an audit institution is not fully in control of the relevance of the topics. What was policy relevant at the time of the selection may no longer be

relevant when the report is published. Chance events can also hinder or boost the impact.

- Drawing on lessons from the literature dealing with policy evaluation utilization, SAIs need to make sufficient effort to disseminate their audit results and confer with all relevant stakeholders.

Based on our research, lessons can also be learned for academic research on performance audit. More and better impact research is necessary if we are to come to more relevant and reliable conclusions on the degree and nature of the impact as well as the factors facilitating impact. In such research it would be highly desirable to examine a wide range of impact dimensions, and particular attention should go to factors at the micro-, meso- and macro level and their mutual interactions. In particular, attention should be focused on SAIs that have not been included in the literature yet; on comparative research that takes into account institutional and other differences between the SAIs; on the way the dilemma between SAIs' watchdog function and counsellor function influences the impact of the audits; and on barely explored forms of impacts such as anticipative impact and political use of audits.

NOTES

1. An extended version of this chapter was published earlier in a Dutch language book published by the Flemish Government (*Studiedienst van de Vlaamse Regering, Efficiëntie en effectiviteit van de publieke sector in de weegschaal*, 2010).
2. Johnsen, A., Meklin, P. (2001); Schelker, M., Eichenberger, R. (2008); Torgler, B., Schaltegger, C. A. (2005); Weets, K. (2008).
3. For instance, two of the four studies on local audit bodies cover the Swiss local finance commissions. They can hardly be compared with national SAIs. In some districts these commissions are even empowered to audit *ex ante* policy proposals and suggest other alternatives.
4. Extensive research is done on the factors that promote or hinder the use of scientific knowledge. A great deal of these studies highlight the importance of matches or mismatches between the scientific community and the community of policymakers and managers. A three-community theory emphasizes the importance of so-called advocacy coalitions, coalitions of persons from different organizations who share a same set of normative and causal persuasions and often act jointly to influence a policy. Another range of factors mentioned are the characteristics of the decision-making process which influence the use of scientific knowledge. Characteristics of the organization to which the evaluation or the research relates and characteristics of the relevant decision-makers are not so often cited as explaining factors.
5. One study relates to data dating back to the seventies, three studies to the eighties, two studies to the eighties and nineties, one study to the nineties, five to the nineties and to the years 2000 and two to the year 2000.
6. The United States (4), Canada (2), Israel (2), the Netherlands (2), Australia (1), Estonia (1), United Kingdom (1), set of countries (1).
7. European Court of Auditors (2008); Azuma (2004); Talbot (2010).

8. UN / INTOSAI (2007).
9. The Working Group on Value and Benefits of SAIs (http://www.intosai.org/en/portal/committees_working_groups_task_forces/goal_3_knowledge_sharing/wgvalue/benefitssais/).
10. European Court of Auditors (2008).
11. For instance Azuma (2004). For the Australian National Audit Office there is 91% of recommendations agreed by audited entities. For the GAO the 'percentage of recommendations made by the GAO four fiscal years ago (in fiscal 1998) that have been followed by audited entities' was 72%.
12. Dolmans and Leeuw (1997); Elliott (2002); Frey and Serna (1990); Frey (1994); Leeuw (2000); Leeuw (2006); Munro (2004); Power (1999); Power (2003); Van Thiel and Leeuw (2002).
13. For example: an independent variable can be a condition, necessary or sufficient or sufficient but not necessary or else simply a contributing factor in addition to many other factors; causality can be linear or circular; there can also be a cumulative effect of several audits (Weets, 2008).

REFERENCES

- Algemene Rekenkamer (2007), *Report 2007*, The Hague: Algemene Rekenkamer.
- Alon, G. (2007), 'State audit and the media as the watchdogs of democracy – a comparative view', *Lyunim – The Periodical of the Office of the State Comptroller and Ombudsman*, **61**, 55–100.
- Australian National Audit Office (2008), *Annual Report 2007–2008*, Commonwealth of Australia, 2008, available at: (<http://www.anao.gov.au/director/aboutus/prevannualreport.cfm> (accessed 13 May 2010)).
- Azuma, N. (2004), 'Performance measurement of Supreme Audit Institutions' *Government Auditing Review*, Board of Audit of Japan, March, 65–94.
- Bemelmans-Vidéc, M.L., Lonsdale, J. and Perrin, B. (2007), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, Transaction Publishers, US, New Jersey, 296 p.
- Blume, L. and Voigt, S. (2007), 'Supreme audit institutions: supremely superfluous? – A cross country assessment, International Center for Economic Research', Working Paper 3/2007.
- Bongers, F. (2009), 'Preventief gedrag bij beleidsevaluaties' (preventive behaviour in policy evaluations), in Verlet, D. and Devos, C., *Over Beleidsevaluatie: Van Theorie naar Praktijk en Terug*, Studiedienst van de Vlaamse regering, Brussels, 62–76.
- Brans, M. and Van Der Straeten, T. (2004), 'Het gebruik van kennis: naar een geïntegreerd model van interacties tussen wetenschap, beleid en maatschappij' (the use of knowledge: towards an integrated model of interactions between science, policy and society) in Brans, M., Jacobs, D., Martinello, M., Rea, A., Swyngedouw, M., Adam, I., Balancier, P., Florence, E. and Van Der Straeten, T. *Onderzoek en beleid: de gevalstudie van immigratie in België* (research and policy: immigration in Belgium – case study), pp. 13–31.
- Brown, R. (1980), 'Auditing and public administration: the unrealized partnership', *Public Administration Review*, May/June, 259–65.
- Campbell, S., Benita, S., Coates, E., Davies, P. and Penn, G. (2007), *Analyses for Policy: Evidence-Based Policy in Practice*, Government Social Research Unit, London: HM Treasury.

- De Vries, G.J. (2000), *Beleidsdynamica als sociale constructie. Een onderzoek naar de doorwerking van beleidsevaluatie en beleidsadvisering* (policy dynamics as social construction. An investigation into the permeation of policy evaluation and policy counselling), Delft: Eburon.
- Dolmans, L.J.F., and Leeuw, F.L. (1997), 'Performance auditing. Vraagtekens bij de effectiviteit van een onderzoekstraditie', *Naar een doelmatiger overheid. Goed georganiseerd mensenwerk* (Towards a more efficient government. Well-organized people's work), NIVRA pp. 81–8.
- Elliott, J. (2002), 'The impact of intensive "value for money" performance auditing in educational systems', *Educational Action Research*, **10** (3), 499–506.
- Etverk, J. (2002), 'Measuring performance audit effectiveness: the case of Estonia', Master's Thesis, University of Tartu.
- European Court of Auditors (2008), *Report on Preliminary Analysis of Feedback from the External Survey (SAI's) on Performance Indicators*, Luxembourg: European Court of Auditors (unpublished).
- Frey, B.S. (1994), 'Supreme Audit Institutions: a politico-economic analysis', *European Journal of Law and Economics*, **1**, 169–76.
- Frey, B.S., and Serna, A. (1990), 'Eine politisch-ökonomische Betrachtung des Rechnungshofs', *Finanzarchiv*, **48** (2), 244–70.
- Hanney, S., Gonzales-Block, M., Buxton, M. and Kogan, M. (2003), The utilization of health research-policy-making, concepts, examples and methods of assessment, in *Health Research Policy and System*, available at: <http://www.health-policy-systems.com/content/1/1/2> (accessed 13 May 2010).
- Hanney, S., Gonzalez-Block, M., Buxton, M. and Kogan, M. (2003), 'The utilization of health research-policy-making, concepts, examples and methods of assessment', *Health Research Policy and System*, **1** (2).
- Hatherly, D.J. and Parker, L.D. (1988), 'Performance auditing outcomes: a comparative study', *Financial Accountability and Management*, **4** (1).
- Haynes Watts (2004), 'National Audit Office VFM Audit 2003/4. Examination of the NAO's financial impacts', Haines Watts, May 2004. Audit report.
- Ingraham, P.W. and Moynihan, D.P. (2001), 'When does performance information contribute to performance information use? Putting the factors in place', Campbell Public Affairs Institute Working Paper.
- Innvaer, S., Vist, G., Trommald, M. and Oxman, A. (2002), 'Health policy-makers' perceptions of their use of evidence: a systematic review', *Journal of Health Services, Research and Policy*, **7**(4), 239–44.
- INTOSAI (2007), *Building Capacity in SAIs. A Guide*, INTOSAI Capacity Building Committee, November 2007, available at: [http://cbc.courdescomptes.ma/index.php?id=20&tx_abdownloads_pi1\[action\]=getviewcategory&tx_abdownloads_pi1\[category_uid\]=20&tx_abdownloads_pi1\[cid\]=81&cHash=bf236a4c1a](http://cbc.courdescomptes.ma/index.php?id=20&tx_abdownloads_pi1[action]=getviewcategory&tx_abdownloads_pi1[category_uid]=20&tx_abdownloads_pi1[cid]=81&cHash=bf236a4c1a) (accessed 13 May 2010).
- Johnsen, A. and Meklin, P. (2001), 'Performance auditing in local government: an exploratory study of perceived efficiency of municipal value for money auditing in Finland and Norway', *The European Accounting Review*, **10** (3), 583–99.
- Johnson, K., Greenseid, L.O., Toali, S.A., King, J.A., Lawrenz, F. and Volkov, B. (2009), 'Research on evaluation use. A review of the empirical literature. From 1986 to 2005', *American Journal of Evaluation*, **30** (3), 377–410.
- Johnston, W.P. (1988), 'Increasing evaluation use: some observations based on the results of the U.S. G.A.O.', *New Directions for Program Evaluation*, **39**, 75–84.

- Kirkhart, K.E. (2000), 'Reconceptualizing evaluation use: an integrated theory of influence', *New Directions for Evaluation*, **88**, 5–23.
- Lapsley, I. and Pong, C.K.M. (2000), 'Modernization versus problematization: value-for-money audits in public services', *The European Accounting Review*, **9** (4), 541–67.
- Leeuw, F.L. (2000), 'Unintended side effects of auditing: the relationship between performance auditing and performance improvement and the role of trust', in Raub, W. and Weesie, J. (eds.), *Management of Durable Relationships. Theoretical Models and Empirical Studies of Households and Organizations*. Proceedings of the Conference The Management of Durable Relations, Thesis Press, Amsterdam (2000).
- Leeuw, F.L. (2006), 'Over impact en neveneffecten van performance monitoring en auditing', *Vlaams Tijdschrift voor Overheidsmanagement*, 36–44.
- Leviton, L. (2003), 'Evaluation use: advances, challenges and applications', *American Journal of Evaluation*, **24** (4), 525–35.
- Lindquist, E.A. (1990), 'The third community, policy inquiry and social scientists', in Brooks, S. and Gagnon, A.-G. (eds.), *Social Scientists, Policy and the State*, New York: Praeger, pp. 21–51.
- Milgrom, T. and Schwartz, R. (2008), 'Israel's auditor as policy change agent: the case of water policy', *International Journal of Public Administration*, **31** (8), 862–77.
- Morin, D. (2001), 'Influence of value for money audit on public administrations: looking beyond appearances', *Financial Accountability and Management*, **17** (2), 99–118.
- Morin, D. (2003), 'Controllers or catalysts for change and improvement: would the real value for money auditors please stand up?', *Managerial Auditing Journal*, **18** (1), 19–30.
- Morin, D. (2004), 'Measuring the impact of value-for-money audits: a model for surveying audited managers', *Canadian Public Administration*, **47** (2), 141–64.
- Morin, D. (2008), 'Auditors general's universe revisited. An exploratory study of the influence they exert on public administration through their value for money audits', *Managerial Auditing Journal*, **23** (7), 697–720.
- Munro, E. (2004), 'The impact of audit on social work practice', *British Journal of Social Work*, **34**, 1075–95.
- National Audit Office (2008), *Annual Report 2008*, London: National Audit Office.
- Nutley, S., Davies, H. and Walter, I. (2002), 'Evidence Based Policy and Practice: Cross Sector Lessons From the UK', ESRC UK Centre for Evidence Based Policy and Practice, Working Paper 9.
- Nutley, S., Webb, J. (2000), 'Evidence and the policy process', in Davies, H., Nutley, S. and Smith, P. (eds.), *What Works? Evidence-Based Policy and Practice in Public Services*, Bristol, UK: The Policy Press, 31–41.
- Office of the Auditor-General of Canada (2006–07), *Estimates Report on Plans and Priorities*, Office of the Auditor-General of Canada, 2006, available at: <http://www.tbs-sct.gc.ca/rpp/2006-2007/OAG-BVG/OAG-BVG00-eng.asp> (accessed 13 May 2010).
- Oh, C.H. (1997), 'Explaining the impact of policy information on policy making, knowledge and policy', *The International Journal of Knowledge Transfer and Utilization*, **10** (3), 25–55.

- Patton, M.Q. (1997), *Utilization-Focused Evaluation: The New Century Text*, Thousand Oaks, California: Sage.
- Power, M. (1999), *The Audit Society. Rituals of Verification*, Oxford: Oxford University Press.
- Power, M. (2003), 'Evaluating the audit explosion', *Law & Policy*, **25** (3), 185–202.
- Put, V. (2006), *De bril waarmee auditors naar de werkelijkheid kijken. Over normen die rekenhoven gebruiken bij het beoordelen van de overheid*, Brugge: die Keure.
- Sadcliff, V.S. (1997), 'Competing rationalities in "special" government audits', *Critical Perspectives on Accounting*, **8** (4), 343–66.
- Rossi, P.F., Freeman, H.L. and Lipsey, M.W. (1999), *Evaluation: A Systematic Approach*, Thousand Oaks, California: Sage.
- Saint-Martin, D. (2004), 'Managerialist advocate or "control freak"', The Janus faced Office of the Auditor General', *Canadian Public Administration*, **47** (2), 121–40.
- Schelker, M. (2007), 'Public auditors: empirical evidence from the US states', available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=959392 (accessed 2 August 2010).
- Schelker, M. and Eichenberger, R. (2008), 'Rethinking public auditing institutions: empirical evidence from Swiss municipalities', available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1028927 (accessed 13 May 2010).
- Sutcliffe, S. and Court, J. (2005), *Evidence-Based Policymaking: What Is It? How Does It Work? What Relevance for Developing Countries?* Overseas Development Institute (ODI), November 2005, <http://www.odi.org.uk> (accessed 20 April 2010).
- Talbot, C. and Wiggan, J. (2010), 'The public value of the National Audit Office', *International Journal of Public Sector Management*, **23** (1), 54–70.
- Torgler, B. and Schaltegger, C.A. (2005), 'The determinants of political discussion: how important are audit courts and local autonomy?' CREMA, working paper 2005–28.
- UN/INTOSAI (2007), 'The value and benefits of audit in a globalised environment, Report on the 19th UN/INTOSAI Symposium on Government Audit, Vienna, 28 to 30 March 2007, UN (Division for Public Administration and Development Management, Department of Economic and Social Affairs (DESA)) and INTOSAI, 23 p. + annexes (140 p.) http://www.intosai.org/en/portal/events/un_intosai_seminars/chronology/ (accessed 13 April 2010).
- Van Der Meer, F.B. (1999), 'Evaluation and the social construction of impacts', *Evaluation*, **5** (4), 387–406.
- Van Thiel, S. and Leeuw, F.L. (2002), 'The performance paradox', *Public Performance and Management Review*, **25** (3), 267–81.
- Varone, F., Jacob, S. and De Winter, L. (2005), 'Polity, politics and policy evaluation in Belgium', *Evaluation*, **11** (3), 253–73.
- Weets, K. (2008), 'How effective are performance audits? A multiple case study within the local audit body of Rotterdam', Paper presented at the 5th International Conference on Accounting, Auditing & Management in Public Sector Reforms, EIASM, Amsterdam, 3–5 September 2008.
- Weiss, C. (1979), 'The many meanings of research utilization', *Public Administration Review*, September/October 1979, 426–31.
- Weiss, C. and Bucuvalas, M. (1980), 'Truth tests and utility tests: decision-makers'

frames of reference for science research', *American Sociological Review*, **45**, 302–13.

World Bank (2003), *Survey on Budget Practices and Procedures*, available at: <http://www.oecd.org/dataoecd/14/18/36930865.pdf> (accessed 20 April 2010).

10. Accountability, performance and performance auditing: reconciling the views of scholars and auditors

Mark Funkhouser

Drawing on the limited empirical evidence available, Van Loocke and Put concluded that performance audit can and does have impact, although it varies considerably from project to project. This view tallies with that of many performance auditors, who believe that performance auditing can play a critical role in effective democratic governance by strengthening accountability and improving performance. However, these beliefs seem to be flatly contradicted by the work of a number of academics such as Behn (see Behn, 2001) and Dubnick (see Dubnick, 2005), two of the most well-respected scholars in the field of accountability and public management. This chapter aims to examine, understand and reconcile the differences between the views on accountability, performance, and performance auditing held by Behn, Dubnick and other scholars and the views held by many of the auditors.

THE BANE OF PERFORMANCE AUDITING

As has been observed elsewhere (Funkhouser, 2008), despite its connections with ‘new public management’, leading public management thinkers and practitioners have generally ignored performance audit. Where they have commented, their views have tended to be negative and they have ‘employed a range of arguments, from democratic theory to extensive empirical evidence, to persuade us that performance auditing’s impact on public organizations is at best benign, and at worst malevolent’ (Funkhouser, 2008).

In *Rethinking Democratic Accountability* (2001) Robert Behn offers a scathing denunciation of performance auditing. In a section entitled ‘The Bane of Performance Auditing,’ he writes:

The problem with performance *auditing* is that auditors believe that theirs is the only profession that is capable, authorized, or entitled to do it. And if the auditors manage to convert their version of ‘performance auditing’ into the primary vehicle for creating accountability for performance, performance accountability will become compliance accountability. If this happens, accountability will come to mean nothing more than whether the agency – the performance auditee – has complied with the performance auditor’s definition of performance. And for the performance auditor to do its audit, it will have to establish some rules, regulations, standards or other criteria so that it can audit the agency’s records and behavior to determine whether it has indeed complied with the performance criteria. (Behn, 2001, 203)

Behn continues by arguing that whatever areas auditors expand into they will always audit, modifying their practices only slightly, with their focus still on compliance. In addition, Behn has at least three other major criticisms of ‘institutions of accountability’, of which auditors and performance auditing are a part. He argues that they are nit-picking, often pouncing on small errors, hierarchical and punishing, and ultimately they undermine the public trust needed to allow government to function well. As a result, ‘we ought not to even use the phrase “performance audit,” lest we give the auditors the excuse to assert a unique, professional claim to the task of creating performance accountability’. (Behn, 2001, 202–03)

ACCOUNTABILITY IMPROVES PERFORMANCE?

In ‘Accountability and the Promise of Performance: In Search of the Mechanisms,’ Melvin Dubnick carefully examines the relationship between accountability and performance, using a ‘social mechanisms’ approach to determine how and why increased accountability could lead to increased or improved organizational performance. He concludes that ‘This search for mechanisms indicates that the relationship is paradoxical and either spurious or so contingent as to raise questions regarding the administrative reforms based on it’ (Dubnick, 2005, 376).

Along with Behn and other scholars (Michael, 2005; Frederickson, 2005) Dubnick makes the point that the concept of accountability is used widely and loosely in contemporary discussions of government performance. He argues that the idea that accountability increases performance has been accepted without careful examination and that those who have carefully studied the relationship have discovered an ‘accountability paradox’, in which more accountability actually diminished or impeded organizational performance.

This situation helps explain a phenomenon that is best termed the ‘accountability paradox.’ Halachmi, Bouckaert, and others (Bouckaert & Peters, 2002; Halachmi, 2002a; Halachmi, 2002b; Halachmi & Bouckaert, 1994) have pointed out that there exists an inherent tension between accountability and performance on an operational level as well as logically: Increasing efforts to improve performance through accountability tends to have the opposite effect. Rather than acting as a driver for desired levels of improved performance, accountability tends to be a ‘breaker’ by either slowing down or stopping improvements. (Dubnick, 2005, 396)

While casting much doubt on the argument that accountability increases performance, and strongly suggesting that perhaps the opposite might be true, Dubnick still leaves the door open for the possibility that the relationship may hold.

The central argument of this paper is that we cannot – and should not – continue to rely on the assumed relationship between accountability and performance that underlies much of the NPM reform agenda. This position is not based on the contention that the assumption is false, but rather that it is unarticulated and untested. (Dubnick, 2005, 402)

Most significant, we found nothing in the existing literature on either topic that would provide a logical (let alone a theoretical or empirical) link between account giving and performance (outside the performative nature of account giving itself). What emerges, instead, is some insight into the accountability paradox and the awareness that we need to broaden our perspective if we are to find the M factor(s) that might link account giving and performance. If such links exist (and that remains an open empirical question), then they are probably exogenous to the relationship itself. (Dubnick, 2005, 403)

AUDITORS INTEND FOR THEIR WORK TO IMPROVE PERFORMANCE

Government auditors, especially those doing a significant amount of performance auditing, intend for their work to improve the performance of the agencies they audit. By ‘improved performance,’ they mean much more than compliance with rules or following the proper process and procedure. They expect improvement in the end results – the outcomes – of government operations. Auditors also believe that the performance improvement they seek will occur as a result of improved accountability, a point that Jan-Eric Furubo explored in more detail in Chapter 2. This assumed relationship is the fundamental rationale for performance auditing and is articulated in the public statements of the heads of audit agencies, professional articles written by auditors, and in the mission statements of many audit agencies. The quotes below are meant to illustrate both the diversity

of types of organizations and the similarity of the themes articulated in these statements.

The U.S. Government Accountability Office (GAO) is the supreme government audit organization in the United States. GAO promulgates the *Government Audit Standards*, which constitute the core of authoritative literature regarding government audit practice in the United States. Audit organizations auditing government agencies, programs and activities are nearly always required to follow these standards and the standards play a substantial role in shaping audit practice in the United States and throughout the world. David Walker, former Comptroller General of the United States and head of the GAO, said 'We're in the business of helping government work better and holding it accountable to the American people. To this end, GAO provides Congress with oversight of agency operations, insight into ways to improve government and foresight about future challenges' (Walker, 2005).

The Office of the Auditor General of Canada discusses the role of performance auditing as one type of legislative auditing:

The Auditor General's performance audit reports contain recommendations that can serve as a springboard to lasting and positive change in the way government functions. Follow-up audits may be conducted to determine whether the government has made satisfactory progress in implementing the Office's recommendations. (Office of the Auditor General of Canada, 2010)

The Office of Inspector General of the U.S. Environmental Protection Agency (EPA) perform audits of the EPA and its contractors, amongst other things to promote economy and efficiency, and to prevent and detect fraud, waste, and abuse. This is done to serve as a catalyst for improving the environment by helping the Agency operate more economically, effectively, and efficiently (Office of the Inspector General, US Environmental Protection Agency, 2010).

The Legislative Division of Post Audit, the audit organization of the state of Kansas, contracts with a commercial audit firm for the financial audit of the state and devotes its staff resources almost exclusively to performance auditing. The division was created in 1972 expressly for the purpose of conducting performance audits and is one of the oldest such agencies in the United States. What follows is the description of the agency from its website.

The Legislative Division of Post Audit is a staff agency of the Kansas Legislature. . . . It conducts performance audits to determine the efficiency, effectiveness, economy, or results of governmental agencies or programs, or to provide other specific information about those agencies or programs.

Performance audits provide the Legislature and agency management with independent analysis, evaluation, and recommendations regarding the performance of the audited activities. The Legislative Division of Post Audit follows all requirements for conducting performance audits established by the U.S. General Accounting Office [sic]. (Kansas Legislative Division of Post Audit, 2010)

The City of Portland, Oregon was one of the first in the United States to enable its government audit organization to conduct performance audits. The website for the Portland Office of the City Auditor shows the following as the mission of its Audit Services Division.

The mission of the Audit Services Division is to promote honest, efficient, effective, and fully accountable City government. To fulfill this mission the Division audits and evaluates City operations to provide useful and objective information to citizens, City Council, and management. (Portland Office of the City Auditor, 2005)

And all of these statements are broadly in line with *The Implementation Guidelines for Performance Auditing*, published by the International Organization of Supreme Audit Institutions, which defines performance auditing as follows. ‘Performance auditing is an independent examination of the efficiency and effectiveness of government undertakings, programs or organizations, with due regard to economy, and the aim of leading to improvements’ (INTOSAI, 2004, 11).

These articulations of the aims of performance are not restricted to official statements. In interviews conducted in audit offices around the globe in 2006, auditors overwhelmingly defined a successful audit as one that has a positive benefit on the jurisdiction or program being audited. They regularly used the words ‘change’ and ‘impact’, seeing their work as leading, amongst other things, to improvements in conditions in the field they had audited, better processes for delivering services and a more informed debate about performance in an area of government (Funkhouser, 2008). They were united in believing that their work went beyond examining whether there had been compliance with rules.

DO THE PERFORMANCE IMPROVEMENTS THE AUDITORS INTEND ACTUALLY OCCUR?

Auditors believe that their work has sparked actual improvements in performance of government agencies, but at this point you more or less have to take their word for it because – as Van Loocke and Put showed in

Chapter 9 – there has been limited scholarly examination of this question. GAO, for example, publishes annually a detailed analysis of the results of its work. A recent report states that its work led to \$43 billion in financial benefits – ‘a return of \$80 for every dollar invested in us’ and ‘over 1,300 nonfinancial benefits that helped to change laws’ (GAO, 2009, 3). The EPA Office of the Inspector General reports a wide range of accomplishments for the fiscal year 2009. Examples reported include 38 EPA policy, directive, practice or process change/decisions, 103 actions taken or resolved prior to report issuance, 785 recommendations for improvement, \$62.3 million recommended efficiencies, costs saved (Office of the Inspector General, U.S. Environmental Protection Agency, 2010, 3). Many other audit organizations prepare similar reports of the results of their work. For example, the Office of the City Auditor, San Jose, California reports that, on average, for every \$1 of audit costs, ‘we have identified about \$9 in cost savings and/or increased revenues’ and in the period 1985–2003 reported savings and increased revenue of \$182 million compared to audit costs of \$25 million (City of San Jose, 2005, 2010).

In the United Kingdom, the NAO has a target ratio of benefit to cost of £10 for every £1 spent, which one official there described as a ‘litmus test’ on the body’s performance. The organization is very focused on it as a performance measure and an elaborate system for measuring and validating the numbers has been built up to give it credibility. The financial impacts reported by the NAO are agreed with the clients and reviewed by external auditors, who have commended the NAO for the rigor of the approach, the like of which they could find in no other audit office (Funkhouser, 2008).

In addition to reports of overall results of their work, auditors offer anecdotal evidence of performance improvements resulting from specific audits. A request to the members of the National Association of Local Government Auditors in 2005 produced numerous such examples, a few of which are related below.

Example 1

One recent example that comes to mind from our shop relates to our June 2004 Milwaukee County employee/retiree health care benefit audit. We learned in the course of our audit that, although the County is self-funded for approximately 2/3 of its \$112 million (2005 budget) health care plan. . . . As a result of the audit, this particular hospital system more than doubled its discounts (from 10% to 22% – still not up to the market, but an improvement) and another system offered an additional 5% ‘rebate’ on all charges for total savings of roughly \$3 million in 2005. Not huge in light of the overall program cost, but quite a return for simply shedding some public light on a very secretive discount arrangement. As a post-script, the County Executive just released his 2006 budget request yesterday, and it cites projected savings of an additional \$11.5

million from health care plan design changes, almost all of which stemmed from our 2004 audit.

Example 2

We did a performance audit of our EMS (ambulance) service a few years ago during which we discovered that our collection rate was about 37% and annual revenue was about \$750,000. We found that patient data collection contained around 40% error rate – we weren't collecting enough personal data to bill in many cases. We recommended increasing accountability for collecting patient data, and guess what . . . Our collection rate today is 58% and annual revenue is about \$1.5 million.

Example 3

. . . a few months ago we issued a report on our General Obligation Bond Program. We finance bonds through property tax revenue (which can't be used for operations, just debt service). The lion's share of projects financed with GO Bonds are street projects. Our city is consistently rated as one of the ten worst cities in the U.S. in terms of street conditions. The lament for years has been that we don't have enough money to fix the streets. What our audit found was that the city has a lot of money and a great deal more in the pipeline (property values continuing to climb), it just isn't being spent . . . Since our report was released the following has happened: Public Works Director resigned; Engineering Division reorganized; Additional staff have been added including a GO Bond Program Manager; An Executive Steering Committee (City Manager, Assistant City Managers, and several department heads) has been established; Program reporting improved; Spending has doubled and projects are being completed. But most importantly, as a starting point, the City Manager has dictated a performance goal of \$5 million dollars per month (double the historical average of \$30 million per year) in spending. Focusing on simple, irrefutably important measures was absolutely the most effective way of making our entrenched, micromanaging, Public Works Director accountable for his ineffective operation.

Example 4

We disclosed ridiculous instances of overstaffing in the Parking Authority: directors, assistant directors, executive assistants, managers, assistant manager, supervisors, assistant supervisors, work leader, etc. The Parking Authority was the only organization I had ever seen that had a Human Resources Department and a Personnel Department. The staffing was scaled down in response to our audit.

ACADEMIC STUDIES OF PERFORMANCE AUDITING ARE RARE

As Van Loocke and Put noted earlier in this book, there has been little in the way of scholarly work directly examining performance auditing impact. In an article published in the Israeli journal *Studies in State Audit* entitled 'Are the Watchdogs Really Watching? Assessment of Audit Institutions in the Fifty United States of America,' Friedberg and Lutrín (Friedberg and Lutrín, 2004, 115) report conducting a systematic search

for academic articles on public auditing in the United States and finding only two such articles: Wheat's 'The Activist Auditor: A New Player in State and Local Politics' and Baber's 'Another Look at Auditing by State Governments' (Wheat, 1991; Baber, 1995). In addition to these articles, two dissertations, *The Spread of Performance Auditing Among American Cities* (Funkhouser, 2000) and *National Systems for Accountability, Transparency, and Oversight of Economic Activities* (Zacchea, 2001), can be added to that total. To these, we can add a small number of further studies in Canada (for example, Morin, 2001) and others from Europe and Australia, which have been drawn on in this book. Of course, a much larger number of books and articles touch on government auditing as part of a considerably broader discussion of accountability and performance, but it remains the case that the literature is limited.

Only two of the United States studies, those by Funkhouser and by Wheat, focus specifically on performance auditing. This is surprising given the rapid increase in the amount of performance auditing that appears to be occurring and the level of concern that increase has created among some scholars. Funkhouser reported in 2000:

In the literature on diffusion of innovations, an innovation is said to have reached critical mass when 'enough individuals have adopted the innovation so that the innovation's further rate of adoption is self-sustaining' (Rogers 1995, 313). Critical mass occurs when five to twenty five percent of the population has adopted the innovation. Critical mass can be seen on the S-curve as the point of inflection, where the slope of the line becomes steeper. The data for this study show the adoption of performance auditing is past the point of critical mass. The point of inflection of the S-curve, shown in Figure 10.1, occurred in about 1990. Presently, audit directors in exactly half of the cities (109 of 218) report the presence of an audit function doing performance audits. (Funkhouser, 2000, 106)

Wheat's article accurately anticipated the growth in performance auditing and viewed the development with approval. He wrote:

Because of these fiscal, ideological, legal, institutional, and political developments in the public administration context, it is clear that performance auditing is not a transitory fad . . . James P. Wesberry, a past president of the Institute of Internal Auditors, recently suggested that changes in American culture have undermined the concept of accountability and have 'eliminated shame for the act of dishonesty. We have substituted explanations which transfer the blame to others – or to no one.' The result, according to Wesberry, is that the present generation of auditors 'will have to audit the unashamed – those who have no fear of doing wrong, who have no regret for taking more than they deserve, who feel regret only if they have the "bad luck" to be found out' (1989, pp. 25–26). In such an amoral political and administrative culture,

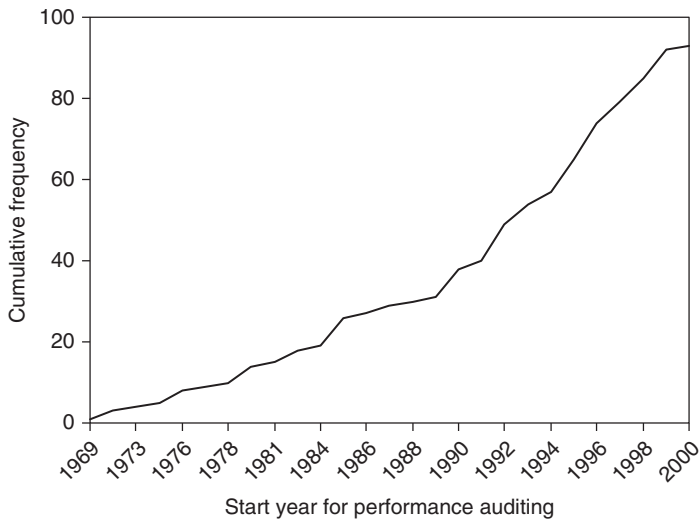


Figure 10.1 *Performance audit adoption by year in cities in the United States with populations of 100,000 or greater*

perhaps we should welcome the emergence of new players in the great game of state and local politics who possess a bias toward disclosure and carry about a little book of publicly acknowledged and professionally agreed upon principles and standards to which they feel obligated to adhere. (Wheat, 1991, 388 and 391)

In their scathing critique of anti-corruption efforts in the United States, Anechiarico and Jacobs also note, with apparent favor, the growth of government auditing and its spread into performance auditing.

Public-sector auditing clearly has a bright future. Auditing serves so many functions for so many interests inside and outside government that its growth in size and influence is certain. We foresee the continued proliferation of auditing entities within and external to operating agencies. There is every reason to believe that auditing will become more intensive and comprehensive as it makes greater use of the new information and monitoring technologies. Auditors themselves will become increasingly sophisticated, influential, and law-enforcement oriented. (Anechiarico and Jacobs, 1996, 192)

Of the control mechanisms we have considered, auditing and accounting have the greatest potential to further, simultaneously, the goals of preventing corruption and of improving the efficiency and effectiveness of public administration. Strong accounting and auditing programs are essential to good management and for identifying and evaluating the costs of operations and services. (Anechiarico and Jacobs, 1996, 201)

Regardless of whether the growth in performance auditing is seen as a good, bad, or mixed development (and Frans Leeuw's chapter presents a more skeptical perspective), it seems very likely that it will continue. A study published in the United States concludes that '...the nation's institutional capacity to perform independent performance auditing is lacking relative to probable future demand' (Klay et al., 2004, 137). In part, this will be because performance audits' main preoccupations – accountability and improvement – remain central to debate about the use of public resources, which as Chapter 1 made clear is ever more pressing in the context of the economic crisis. As Bouckaert and Halligan (2008) have put it, 'it is almost certain that "performance" will remain as a focus of public management and policy.'

If that is to be the case there will need to be more evidence of the value achieved by performance audits. Of those mentioned above, only Zacchea's work contains an examination of the outcomes of government auditing. He conducted a comprehensive examination of 179 supreme audit institutions (the top government audit agency within a national government) globally and developed a list of 21 different factors associated with an effective audit function. Then based on interviews and surveys, he rated the audit organizations on each of the factors. He found that the scores of the audit organizations on the 21 factors correlated positively with broad measures of national economic performance such as real economic growth rates (Zacchea, 2001, 105–110). Finding evidence of this kind applicable to performance audit is key to pressing the case for its value.

EXAMINING THE DIFFERENCES IN VIEWS

Examination of Behn's 2001 critique of performance auditing reveals six primary concerns. It is not possible to offer systematic, independently derived evidence to address these concerns given the absence of rigorous scholarly research on performance auditing, but we examine each in turn.

Firstly, Behn notes the lack of clarity with regards to the definition of performance auditing. From the point of view of a professional auditor, there is a certain amount of merit to this claim. Certainly the profession has struggled to reconcile the vast differences in beliefs and interpretations with regard to what constitutes performance auditing, although there has been considerable convergence over the last decade or so. The study of performance auditing in large American cities referred to earlier chronicles this evolution in the definition of performance auditing in the United States (Funkhouser, 2000).

The chapter by Furubo further explores the definition of performance audit. Currently, the views offered by the INTOSAI document *Implementation Guide for Performance Auditing* seem the most authoritative and accurate statements regarding this issue. It states:

Performance auditing is an independent examination of the efficiency and effectiveness of government undertakings, programs or organizations, with due regard to economy, and the aim of leading to improvements (INTOSAI, 2004, 11). Performance auditing is not a regular audit with formalized opinions, and it does not have its roots in private auditing. It is an independent examination made on a non-recurring basis. It is by nature wide-ranging and open to judgments and interpretations. It must have at its disposal a wide selection of investigative and evaluative methods and operate from a quite different knowledge base to that of traditional auditing. It is not a checklist-based form of auditing. The special feature of performance auditing is due to the variety and complexity of questions relating to its work. (INTOSAI, 2004, 12)

Secondly, Behn seems to assert that auditors want exclusive domain over the issue of performance accountability. Auditors do want those who claim that a particular piece of work is a performance audit to have followed the relevant audit standards in conducting the work. But auditors recognize that their work is only a small piece of a much larger on-going process of accountability for performance. The press, the courts, legislative bodies, individual citizen activists, independent ‘watchdog’ groups and special commissions such as the 9/11 Commission are all part of the process. Often audits are triggered by the work of one or more of these other players. Sometimes audits cause one or more of these players to begin to focus on an issue, as in the case of the government sponsorship program in Canada, mentioned later (Fraser, 2007). There is also considerable evidence from the UK of performance auditors making extensive use of outside experts and consultants to carry out audits with and for them, including whole studies in some cases (Lonsdale, 2008).

Thirdly, Behn argues that performance audit will become compliance auditing in which auditors will seek to determine whether ‘the agency – the performance auditee – has complied with the performance auditor’s definition of performance.’ Auditors recognize that the best work is focused on outcomes – as in the examples from the mission statement and work of the Office of the Inspector General, US Environmental Protection Agency mentioned above. Compliance is an important issue, but often non-compliance discovered by the auditor is only a small part of a larger issue. Many experienced performance auditors are mildly contemptuous of what they regard as ‘mere compliance auditing.’ The criteria by which agency outcomes are judged are usually derived largely from what the agency itself represents as best practice in the relevant field, as Vital Put’s earlier chapter discussed.

Fourthly, Behn accuses performance auditors of nit-picking, of pouncing on small errors at the first hint of blood. No doubt some auditors do that. However, performance auditing is difficult and expensive and an audit agency that focuses on small errors will create enemies without having any major positive impact to show for its expense and effort. Such an audit agency will not thrive and probably will not survive long. Audits are selected sometimes by the audit agency itself and sometimes by direction from the legislative body to which the audit agency reports, although, as Put and Turksema in this book have shown, the relative weight of the different influences can be very different in different audit bodies. In either case, what the best performance audit agencies find themselves addressing are what are referred to in the public administration literature as 'wicked problems.' For example, the state of Kansas was embroiled for several years in a serious, bitter controversy over school funding. The state constitution says the state will fund an 'adequate' education but there is deep disagreement over what constitutes 'adequate' funding. The issue went to the state supreme court and the court threw out the solution agreed to by the legislature and told the legislature to come up with a new formula that would have some rational connection to 'adequate.' The legislature turned to its audit arm, the Legislative Division of Post Audit, to tackle this problem because it was the one agency that seemed to have the technical expertise and was viewed by all parties as fair and credible. These are the problems good performance audit agencies get – big ones.

Fifthly, Behn appears to see performance auditors as the superior party in a linear, hierarchical relationship, with power to punish those subordinates being audited. In fact, the relationship is not a linear, hierarchical one between auditors and those being audited. Audits take place in an arena of multiple conflicting and cooperating parties, in which auditors have no authority over any other party other than access to records, which is frequently contested. Often those whom auditors come into conflict with have substantially more formal authority and informal power than the auditors. Caiden captures this relationship well.

Seeing that the other officeholders, especially those with superior powers and responsibilities, follow the law and use public resources responsibly does not merely require technical competence, it requires considerable moral courage too. Obviously, technical competence is a precondition for effective accounting and auditing. Incompetence is next to useless. When auditors were first employed by rulers it was to provide clear, unbiased, untainted, accurate information on the state of public income and expenditure, to know just how many resources were available and how much money was being raised and where it

went, through whose hands it passed and how honest, reliable, and trustworthy were public money handlers. Such knowledge may not have been shared with the rulers but someone in authority had to have accurate accounts, know of public indebtedness and just how much could be raised from lenders, plan future expenditures, pay off old debts, and identify those who gave good value and were true to their word. Inaccuracies and mistakes could lead to disaster and often did until the state of the art improved with some proficiency. But sometimes just obtaining the correct information, assessing the health of public finances, and telling the rulers unpleasant news demanded the highest moral courage. It still does. (Caiden, 2002, 327)

Auditors do not have the powers to make governments take action or generally to determine any form of punishment. What they can do is make government face up to, prevent or surmount its problems. And they can make officials assess the value of their programs, something they tend not to want to do willingly (Schwartz, 1999). This is some way from the image of a punitive auditor.

And finally, Behn worries that performance auditing will undermine the public's trust in government. Trust in government depends on three things: integrity, competence and transparency. Citizens who trust their government believe it will try to do the right thing and tell them the complete unvarnished truth. They believe their government will usually succeed in what it attempts and will learn from mistakes and improve its performance over time. When mistakes are not acknowledged but are covered up learning cannot occur. Finally, citizens believe these things to the extent that they can see the evidence supporting these beliefs with their own eyes because their government is transparent. Audit organizations intend for their work to contribute to increasing the public's trust in government. For example, the stated purpose for the City Auditor's Office of Kansas City, Missouri, states that: 'Our goal is to conduct audits that answer questions that matter to people outside of City Hall; that enable the city to reduce, avoid, or recover costs; and to alert city officials to potential problems that could undermine the public's trust in city government' (Kansas City, 2010).

The Annual Plan of the Controller and Auditor-General of New Zealand also includes an elaborate strategic plan with inputs, processes and outputs that culminate in one end outcome. It states: 'We will measure our impact/contribution to our end outcome – "Trust in an effective and efficient public sector" – by assessing the extent of improvements over time in Parliament's and the public's perceptions of public sector performance and trustworthiness' (New Zealand Controller and Auditor-General, 2005, 30).

ACCOUNTABILITY AND PERFORMANCE: FURTHERING THE SEARCH FOR MECHANISMS

In their professional lives, auditors observe directly situations in which audits seem to lead to increased accountability and improved performance and they talk among themselves about these events. Casting the stories they tell each other in the language of the academic seems to reveal three themes: public dialogue; organizational learning; and political responsiveness. The important search for the mechanisms by which accountability improves performance initiated by Dubnick (2005) could be furthered by rigorous scholarly examination of these themes, each of which is sketched out briefly below.

Public Dialogue

The stories auditors tell about successful performance audits often show the audit as the opening act in a long series of events involving multiple actors that over time lead to a positive shift in performance. Often the audit revealed some new fact that was not known to the major stakeholders involved, or presented known facts in a new light, which resulted in some shift of the paradigm that stakeholders used to understand the meaning and purpose of the program. An excellent example of this public dialogue can be seen in Canada, where the Office of the Auditor General (OAG) audited the federal government's management of three contracts awarded to a communications agency seven years earlier and produced a report that unexpectedly resulted in a public judicial inquiry promoted by the audit findings, criminal prosecutions and jail terms for a few key players. It led to a Commission of Inquiry into sponsorship and advertising activities, with the audit work coming under intense public scrutiny (Fraser, 2007).

In 'Keeping Public Officials Accountable through Dialogue: Resolving the Accountability Paradox,' Roberts (2002) advocates the use of public dialogue to address 'wicked problems'. She recognizes that accountability is not a single event, and is not unidirectional from one 'account giver' to one 'account receiver' but is, in the political arena, multiple conversations all going on at once back and forth among multiple parties. Public dialogue occurs to the extent that parties actually *attend* to what each other is saying. Roberts writes:

Dialogue is a 'special kind of talk' (Dixon, 1996, 24). According to one of its more renowned practitioners, physicist David Bohm, the word 'dialogue' comes from the Greek *dialogos*: *Logos* means 'the word' or the meaning of the

word, and *dia* means ‘through’ (Bohm, 1985, 1990). This derivation suggests the image of ‘a stream of meaning flowing among us and through us and between us – a flow of meaning in the whole group, out of which will emerge some new understanding, something creative’ (Bohm, 1992, 16). Martin Buber (1970) captures the essence of dialogue well in his classic work *I and Thou*. In the *I–Thou* relationship, each person opens up to the concerns of the other. Both parties reach beyond the limited confines of self to eventually say ‘you and me’ rather than ‘you or me’. (Roberts, 2002, 660)

In *The Spirit of Public Administration*, Frederickson offers the idea of public dialogue as a way to address the politics–administration conundrum in public administration. But, he writes:

Public dialogue between friends and equals is a high standard if applied to the process of administering government and its programs. But if the process of administration is political, as thinkers from Aristotle to Dworkin have argued, then *some means* must be found to make administration nonarbitrary from the standpoint of the public. (Frederickson, 1996, 107 [emphasis added])

Performance auditing can be one of the ‘means’ that enables public dialogue. When done well, performance auditing (as the Canadian example illustrates) can strengthen citizenship and contribute to community dialogue by identifying problems and opportunities and enabling citizens to engage in informed debate and discussion with responsible officials about proposed solutions (Funkhouser, 2000, 2002).

Auditors frequently refer to their work as ‘speaking truth to power.’ In his book of that title, Wildavsky wrote, ‘The highest form of analysis is using the intellect to aid interaction between people’ (Wildavsky, 1987, 17). Good performance auditing meets Wildavsky’s test of aiding interaction among people by looking for significant problems, paying careful attention to analysis and evidence, and providing reports that are clear, convincing and useful public documents.

Organizational Learning

In the stories auditors tell about successful performance audits, the improved performance occurs because, as a result of information brought to light in an audit, different stakeholders involved within or outside the organization learn new information that they are able to put together with information they already have to understand problems and proposed solutions in a new light. For example, in an audit of a transit system the system managers said they were willing to share route information with a neighboring transit system in order to improve mobility options for the transit-dependent but that the other system would never be willing

to reciprocate. Through the audit each side was able to see that the other would, in fact, cooperate and that mutual cooperation would be beneficial to both.

Building on the work of Argyris, Leeuw and Sonnischen have shown that performance auditing and program evaluation can be a key part of organizational learning (Leeuw and Sonnischen, 1994). Argyris conceptualizes two types of learning: single-loop and double-loop. In simple terms, single-loop learning occurs when an existing process is improved and double-loop learning occurs when the process is redesigned. The first brings modest improvements in outcomes while the latter brings dramatic improvements. Successful performance audits – audits with significant performance improvements – result in double-loop learning for the organization. The accountability aspect of double-loop organizational learning is characterized by conflict. For example:

Argyris (1982) defines organizational learning as a ‘process of detecting and correcting error.’ It is a process in which an organization continually attempts to become competent in taking action, while at the same time reflecting on the action it takes to learn from its present and past efforts . . . A more comprehensive and systemic learning process occurs when ‘double-loop learning’ occurs. In this case the assumptions underlying the policies and goals of the organization are questioned, leading to the possibility of securing new and innovative permanent solutions to problems.

A double-loop learning environment encourages the questioning of assumptions and confronting the traditions in an organization that are being advocated. In this learning system, people advocate their views in such a way that would invite confrontation so that their positions might be challenged in a public forum.

Double-loop learning always requires an opposition of ideas for comparison: learning occurs when the underlying policies, norms, theories, and objectives of the organization are questioned.

Candid questioning of an organizational policy can create difficult and complex group dynamics that many persons will actively try to avoid. (Leeuw and Sonnischen, 1994, 3–4)

Auditors have struggled with the dilemma for years, with emphasis swinging back and forth between the two poles. As noted elsewhere, ‘We can see the conflict played out in professional publications like *Internal Auditor* and *The Local Government Auditing Quarterly*. Sometimes one side is up, sometimes the other. The change in The IIA’s definition of internal auditing a few years back represented a swing towards the “program improvement” side of the argument, and more traditional auditors who focused on accountability were seen as “gotcha” auditors. Then, post-Enron, the pendulum began to swing back’ (Funkhouser, 2008).

Political Responsiveness

Often in the auditors' stories of successful audits, the release of an audit report changed the political dynamics of a situation in a way that led to significant improvements in performance, as a result of better decisions by those with governing authority over the program. For example, in an audit of a streetlight system in which the city was leasing the lights from the power company, it became clear that the problem of inadequate and expensive street lighting being experienced by the citizens was not an engineering or a management problem. Politically, the power company held all the cards. Presenting the facts in a clear, compelling and public way began a community dialogue that eventually resulted in voter approval of a major bond issue to purchase the system for a reasonable price and upgrade to a state-of-the-art system with a significantly lower cost per light.

In writing about politics and administration in local government, Nalbandian describes how administrators' roles and responsibilities with regard to politics 'are filtered through a set of enduring democratic values: efficiency, representation, social equity, and individual rights.' He argues that these four values constitute political responsiveness (Nalbandian, 1994). Politics is the process by which stakeholders consider and adopt strategies allocating scarce organizational resources and coping with organizational challenges. In politics, stakeholders contend and negotiate with each other over organizational means and ends and the values they embody. Building on Nalbandian's work, 'good politics' – by which is meant politics which reflects the core democratic values – involves a fair and open process in which various stakeholders have relatively equal access to critical information and adhere to the rules and procedures embodied in the governance structure. In good politics, basic values like efficiency, representativeness, social equity and individual rights are legitimated. In 'bad' politics, the process is unfairly tilted in favor of one group of stakeholders, important information is distorted or concealed, and one or more of the core values of community life is abused or ignored. Successful performance audits can increase political responsiveness by 'improving' the politics of a situation to bring about improved performance (Funkhouser, 2003).

CONCLUSION: COMMON GROUND AND A RESEARCH AGENDA

This chapter has considered a complex and tricky issue on which, it is evident here and elsewhere in this book, there are strongly divergent

views. There is some common ground, however. Firstly, it is important to recognize that few scholars or auditors would argue that public performance and accountability are currently fine just as they are. Very few of the auditors from around the world interviewed for the review of performance audit mentioned throughout this chapter (Funkhouser, 2008) could be described as satisfied with the level of impact achieved by their work, and most of the audit organizations had active programs of one sort or another to increase their impact. The works of Behn and of Dubnick cited in this paper are reactions *against* the dominant themes that they see in discussions of accountability and organizational performance. There is no sense of complacency, however, amongst audit practitioners.

In the United States, a brief look at the state of the nation, at least by an auditor, reveals the following.

- The financial condition of the United States government is unsustainable. The former Comptroller General of the United States, David Walker, referred to the nation's worsening financial condition and growing fiscal imbalance as the 'most urgent' challenge facing the country (Walker, 2005, 2), and things have only got worse since then.
- Public administrators, scholars such as Behn, and many auditors believe that trust is essential to effective organizational performance and that trust legitimizes government. But trust in government in the United States has been declining for decades and work by Marlowe shows that citizens' trust in government is strongly correlated with their belief in whether 'the democratic system is working.' A decline in trust then, represents a decline in confidence in the democratic system itself (Marlowe, 2004).
- Basic infrastructure is one of the most critical components of effective delivery of public services. According to the American Society of Civil Engineers, the nation's infrastructure has declined dramatically over the past 17 years. Across 15 different categories of infrastructure, the highest grade issued in their 2005 review was one C. Most grades were Ds. 'In short, U.S. roads, bridges, sewers, and dams are crumbling and need a \$1.6 trillion overhaul, but prospects for improvement are grim' (Kemp, 2005, 78).

Behn, Dubnick and most auditors appear to have at least three main areas of common ground. First, most auditors would agree with Behn on the issue of mutual accountability with regard to organizational performance and problems. In their hearts at least, most auditors know that in fulfilling their roles, Behn has it exactly right when he says:

Any major failure in organizational performance has multiple causes. And many different people connected with the organization – from top leadership and legislative overseers, to frontline workers and stakeholders, to auditors and journalists, to citizens – understand some (if not all) of these causes. They ought to have a positive obligation to identify and help fix them. (Behn, 2001, 70)

Second, Behn refers repeatedly to the ‘necessity of trust’ for a business, an organization, a society or a government to function. As has been shown earlier in this chapter, and is evident in reviewing articles in the professional journals of auditors, most performance auditors agree with this proposition and intend their work to increase public trust and confidence in public institutions.

Finally, auditors should welcome Dubnick’s desire to undertake an ‘analysis that will provide us with at least the logical foundations (i.e. propositions) for a testable argument about the relationship [between performance and accountability]’ (Dubnick, 2005, 403). Auditors should welcome the search for an empirical and theoretical basis for views that they presently hold only as tenets of cherished faith.

An agenda for scholarly examination of performance auditing could include:

- Content analysis of actual audit reports. For example, an analysis of a sample of the recommendations issued by a particular audit organization like GAO or a group of audit organizations like those belonging to the National Association of Local Government Auditors.
- Examination of the impact of specific audit organizations on the governmental jurisdictions to which they belong in terms of performance and policy outcomes.
- Examination of the differences in processes, products and results in audit organizations that arise from structural and mandate differences in the organizations themselves.

Performance auditing is rapidly expanding. Given the nature and importance of the role, it would be better for the auditors and those they seek to serve if development of the profession were grounded in rigorous scholarly theory and empirical research.

REFERENCES

- Anechiarico, Frank and James B. Jacobs (1996), *The Pursuit of Absolute Integrity: How Corruption Control Makes Government Ineffective*, Chicago: The University of Chicago Press.

- Argyris, Chris (1982), *Reasoning, Learning and Action*, San Francisco: Jossey-Bass Publishers.
- Baber, W.R. (1995), 'Another look at auditing by state governments', *Journal of Public Budgeting and Financial Management*, **2**, 22–41.
- Behn, Robert (2001), *Rethinking Democratic Accountability*, Washington: Brookings Institution Press.
- Bohm, David (1985), *Unfolding Meaning: A Weekend of Dialogue with David Bohm*, New York: Ark Paperbacks.
- Bohm, David (1990), *On Dialogue*, Cambridge, MA: Pegasus Communications.
- Bohm, D. (1992), 'On dialogue', *Noetic Science Review*, 16–18.
- Bouckaert, Gert. and John Halligan (2008), *Managing Performance: International Comparisons*, London: Routledge.
- Bouckaert, Gert. and B. Guy Peters (2002), 'Performance measurement and management: the achilles heel in administrative modernization', *Public Performance and Management Review*, **25** (4), 359–362.
- Buber, Martin (1970), *I and Thou*, Trans. by William Kaufmann, New York: Charles Scribner's Sons.
- Caiden, Gerald E. (2002), 'Enron, accountancy, and professional ethics', *Public Integrity*, **4** (4), 321–332.
- City of San Jose (2005), website of the Office of the City Auditor of San Jose at www.sanjoseca.gov/auditor/benefits2/htm (accessed 10 September 2005).
- City of San Jose (2010), website of the Office of the City Auditor of San Jose at www.sanjoseca.gov/auditor (accessed 5 April 2010).
- Dixon, Nancy M. (1996), *Perspective on Dialogue: Making Talk Developmental for Individuals and Organizations*, Greensboro, NC: Center for Creative Leadership.
- Dubnick, Melvin (2005), 'Accountability and the promise of performance: in search of mechanisms', *Public Performance and Management Review*, **28** (3), 376–417.
- Fraser, Sheila (2007), 'Public scrutiny puts SAI audit practices to the test', *International Journal of Government Auditing*, January, 1–5.
- Frederickson, H. George (1996), *The Spirit of Public Administration*, San Francisco: Jossey-Bass Publishers.
- Frederickson, H. George (2005), 'Coming to account', *PA Times*, **28** (5).
- Friedberg, Asher and Carl Lutrin (2004), 'Are the watchdogs really watching? Assessment of the audit institutions in the fifty United States of America', *Iyunim, Studies in State Audit*, **60**, 91–119.
- Funkhouser, Mark (2000), *The Spread of Performance Auditing Among American Cities*, Kansas City: University of Missouri.
- Funkhouser, M. (2002), 'Auditing in America's big cities', *Journal of Government Financial Management*, **51** (2), 20–25.
- Funkhouser, M. (2003), 'Governance and auditing', *Local Government Auditing Quarterly*, **17** (1), 3–4, available at <http://www.nalga.org/qrtyly/articles.html>.
- Funkhouser, Mark (2008), *Honest, Competent Government: The Promise of performance Auditing*, Altamonte Springs: The Institute of Internal Auditors Research Foundation.
- GAO (2009), *GAO Performance and Accountability Report 2009*, Washington DC: Government Accountability Office.
- Halachmi, A. (2002a), 'Performance measurement, accountability, and improved performance', *Public Performance and Management Review*, **25** (4), 370–374.
- Halachmi, A. (2002b), 'Performance measurement: a look at some possible dysfunctions', *Work Study*, **51** (5), 230–239.

- Halachmi, A. and Bouckaert, G. (1994), 'Performance measurement, organizational technology and organizational design', *Work Study*, **43** (3), 19–25.
- International Organization of Supreme Audit Institutions (INTOSAI) (2004), *Implementation Guidelines for Performance Auditing*, available at: www.intosai.org/Level3/Guidelines/3_AudStandComm/3_ImplGPerfAud_E.pdf (accessed 2 February 2010).
- Kansas City, Missouri, City Auditor's Office (2010), 'City Auditor's Office: purpose', at www.kcmo.org/CKCMO/Depts/DeptCityAuditorsOffice/index.htm (accessed 5 April 2010).
- Kansas Legislative Division of Post Audit (2010), 'Legislative post audit process', at www.kslegislature.org/postaudit/about_process.shtml (accessed 20 April 2010).
- Kemp, R. (2005), 'America: on the road to ruin?', *Public Works Management and Policy*, **10** (1), 77–82.
- Klay, William Earle, Sam McCall and Curtis E. Baynes (2004), 'Should financial reporting by government encompass performance reporting? Origins and implications of the GFOA-GASB conflict', in *Financial Management Theory in the Public Sector*, Aman Khan and W. Bartley Hildreth (eds.), Praeger.
- Leeuw, Frans L. and Richard Sonnichsen (1994), 'Evaluation and organizational learning: international perspectives', in Frans Leeuw, Ray C. Rist and Richard Sonnichsen (eds), *Can Governments Learn? Comparative Perspectives on Evaluation and Organizational Learning*, Transaction Publishers.
- Lonsdale, J. (2008), 'Balancing independence and responsiveness', *Evaluation*, **14** (2), 227–248.
- Marlowe, J. (2004), 'Part of the solution, or cogs in the system? The origins and consequences of trust in public administrators', *Public Integrity*, **6** (2), 93–113.
- Michael, Bryane (2005), 'Questioning public sector accountability', *Public Integrity*, **7** (2), 95–109.
- Morin, D. (2001), 'Influence of value for money audit on public administrations: looking beyond appearances', *Financial Accountability and Management*, **17** (2), 99–118.
- Nalbandian, J. (1994), 'Reflections of a "pracademic" on the logic of politics and administration', *Public Administration Review*, **54**, 531–536.
- New Zealand, Controller and Auditor-General (2005), *Annual Plan 2005–06*, available at www.oag.govt.nz/reports/by-type/corporate-documents (accessed 5 April 2010).
- Office of the Auditor General of Canada (2010), 'Performance audits', available at www.oag-bvg.gc.ca/internet/English/au_fs_e_9365.html (accessed 20 April 2010).
- Office of Inspector General, U.S. Environmental Protection Agency (2010), 'About the EPA Office of the Inspector General', available at www.epa.gov/oig/about_epa_oig.htm (accessed 20 April 2010).
- Portland Office of the City Auditor (2010), 'Mission and authority of the Audit Services Division', available at www.portlandonline.com/auditor/index.cfm?c=27097 (accessed 20 April 2010).
- Roberts, Nancy C. (2002), 'Keeping public officials accountable through dialogue: resolving the accountability paradox', *Public Administration Review*, **62** (6), 658–669.
- Schwartz, R. (1999), 'Coping with the effectiveness dilemma: strategies adopted by state auditors', *International Review of the Administrative Sciences*, **65** (4), 511–526.

- Walker, David M. (2005), 'Transforming government to thrive, not just survive, in changing times', Address to the Diamond Cluster Exchange, Kiawah Island, South Carolina, September 19, 2005.
- Wheat, E. (1991), 'The activist auditor: a new player in state and local politics', *Public Administration Review*, **51** (5), 385–392.
- Wildavsky, Aaron (1987), *Speaking Truth to Power: The Art and Craft of Policy Analysis*, 2nd ed., Transaction Publishers.
- Zacchea, Nicholas (2001), *National Systems for Accountability, Transparency and Oversight of Economic Activities*, International Independent University of Muldovia.

11. On the effects, lack of effects and perverse effects of performance audit

Frans L. Leeuw

INTRODUCTION

Van Loocke and Put (Chapter 9) and Funkhouser (Chapter 10) have considered the evidence that performance audit has an impact. This chapter provides comments on the impact of performance audit and, in particular, looks further at one aspect of impact mentioned towards the end of Chapter 9: the evidence of side-effects and perverse effects. As seen elsewhere in this book, during the past 25 years there has been a sharp increase in the attention that society has paid to the effectiveness and efficiency of government organizations and their policies. Important instruments used for this have been cost-benefit analysis, evaluations, performance audits and performance measurements, as well as inspections, reviews and oversight activities. The objective of using these instruments has been not just to gain insight into the efficiency and effectiveness of government policies (and organizations), but also to contribute to the promotion of these goals, as well as to scrutinize, check, regulate and verify (Hood et al., 2004; Gray and Jenkins, 2007).

Lonsdale and Bemelmans-Videc (2007) argue that ‘the growth of scrutiny is a striking feature of late twentieth century bureaucratic life . . . New ways of holding bodies to account are a growth business; there is more information available and more people involved in scrutinising others.’ Power (1997) has described what he termed the creation of an ‘Audit Society’, and such developments can also be seen in the increase in the budgets of audit offices (Martin, 2005), the expansion of performance measurement, target setting and reporting practices (Wilkinson, 2005), and the extension of audit bodies into a growing range of areas such as issuing good practice guidance and reporting on the credibility of government data (Wilkins and Lonsdale, 2007).¹ The increase not only encompasses ‘traditional’ audits, but also new forms that have penetrated many

other aspects of organizations (for example, clinical audit) and in increasingly diverse parts of society. For example, it was reported that the UK Equality Commission envisaged compulsory inspections of companies in a bid to reduce the gap in wages earned by male and female workers – a so-called ‘gender audit’ (Prosser, 2005). In the Netherlands, it has been suggested (although not implemented) that the language that central government uses in official documents should be audited in order to find out to what extent it is gender-neutral.

Given this expansion of the use of ‘audit’ into different areas of public life, it is not surprising that both administrators and politicians can at times have high expectations of performance audits. As Williams (2003) has put it: ‘A good audit does not just follow a money trail and make a de facto report. It evaluates the economy, efficiency and effectiveness of a program so service can be improved. It triggers immediate consequences for any wrongdoing or mismanagement uncovered, and it acknowledges and rewards good performance . . . [W]eak audits [on the other hand,] will damage voter trust’.

The history of performance audit does indeed show how important this work has been and still is. As has been noted recently (Leeuw, 2009), the Dutch SAI, the Algemene Rekenkamer, was one of the driving forces behind the birth and success of innovative public sector performance audits and evaluations in the Netherlands. During the mid-1980s the organization introduced government-wide (that is, comparative) performance audits, that reviewed what ministries and agencies knew about the legality, efficiency and effectiveness of the different tools that were implemented such as subsidies, levies, inspections, public information campaigns, laws and regulations. Elsewhere, the significance of performance audit as the basis for parliamentary scrutiny is also clear. In response to a report by the United Kingdom House of Commons Committee of Public Accounts, drawing on 10 years’ worth of lessons from its examination of NAO audit reports, the UK Government stated that it ‘takes the Committee’s recommendations seriously as the fruit of the accountability process. The best proof of this is that, as the report acknowledges, the great majority of the Committee’s recommendations have been acted upon. The Committee has thus helped the Government to secure financial savings, raise the standards of public services and improve the quality of delivery’ (Treasury, 2006).

This positive perspective on the impact of performance audit is shared and was elaborated on by Funkhouser in this volume. Yet, this is only one side of the picture. Although the establishment of failings regarding efficiency and effectiveness, and the remedying of such failings through performance audit, can contribute to an increase in efficiency and effectiveness, this is not necessarily always the case in practice. Indeed, much social science

literature of the last decade or more has argued that such proceedings may just as easily generate an ‘audit society’, create a ‘performance paradox’, and lead to an overproduction of reports and studies that, in turn, may result in ‘analysis paralysis’ (cf. Van Thiel and Leeuw, 2002; Schwartz and Mayne, 2005).

In essence, phenomena like the ‘audit society’ and ‘performance paradox’ suggest that, firstly, more performance audits do not always result in a more effective and efficient policy and, secondly, that they can give rise to undesired side-effects. Examples of these side-effects can include a reduced commitment on the part of employees in the studied organizations to ‘go for it’, and what has been termed ‘ossification’ (a fear of innovation) (RMO, 2000). These risks have been recognized for some time; for example, more than a decade ago, Bouckaert and Balk (1999) mentioned pathologies that may accompany performance measurement and audits.

The central consideration of this chapter is to ask what is known about the impact of performance audits on the efficiency and effectiveness of policies and organizations. I will also examine the evidence of the extent to which such audits have unintended and undesired side-effects. Because, as van Looke and Put have already shown in Chapter 9, the extent of empirical evidence on effects and side-effects arising from audit is limited, I will make small scale reference to findings from studies on inspectorates and other supervisory agencies. Thirdly, I will consider the question: what causes the undesired and unintended side-effects of performance auditing?

THE IMPACT OF PERFORMANCE AUDITS

That performance audits are important instruments for achieving accountability and transparency, as well as the promotion of efficiency and effectiveness, is wisdom that we – almost – take for granted. The assumption is that performance audits provide us with knowledge and understanding about ‘government failures’, including inefficiency in processes, camel’s noses,² and fraud and corruption, and also about inaccurate policy theories, malfunctioning management information systems, and discrepancies between objectives and means.

Another assumption is that audits contribute to the improvement of efficiency and effectiveness and, amongst other things, to improved service delivery. Furthermore, it is assumed that among the spin-offs of performance audits are more streamlined processes, improved policy design, and greater insight into cost containment. For example, the Canadian Fraser Institute has shown how much the Canadian public sector suffers from inefficiency, basing its comments on public choice theories and the findings

of the Office of the Auditor General of Canada (Clemens et al., 2005). Parliaments often take up such reports, demanding remedies. As Pollitt (2003) notes, 'When auditors present their (critical) evidence, the tendency (amplified by the mass media) seems to be to intensify the transparency and hierarchical accountability of both ministers and public managers. And this may not be an entirely bad thing. In the not-too-distant past it seemed that ministers and senior public servants were rather easily able to conceal and cover up incompetence and wrong doing' (Pollitt, 2003).³

However, most of these examples deal primarily with ambitions and expectations, which we need to distinguish from actual evidence. After all, performance analysts such as auditors may think they can achieve more than they are capable of achieving in reality, in much the same way that we see such behaviour among policy makers, politicians and policy analysts who want to push their own interventions and programs into the spotlight.⁴

Thus, the question remains: what do we actually know about the effects and side-effects of performance audits? Not much, as it turns out, notwithstanding the analysis presented by Van Loocke and Put. Given the increase in auditing activities and organizations over the last few decades, this is remarkable. Recently, Weets (2008) directly addressed the question of how effective are performance audits, summarizing several studies. For example, she noted:

- Johnsen et al. (2001) examined the extent to which performance audits contributed to performance improvement in Finnish and Norwegian local governments. The process of performance audit appeared to be effective in identifying those organizational areas that were in need of review and, in general, local civil servants and politicians did seem to act upon the auditors' reports and recommendations. Furthermore, the authors stated that the information included in the reports was especially used to develop management systems and to improve the quality of the municipal budget.
- Lapsley and Pong (2000) interviewed Scottish auditors who considered performance audits were useful, particularly with regard to the management of public organizations. The benefits were generally said to be of an operational, rather than of a strategic nature. The auditors referred, for example, to the improvement of systems and processes, the adoption of best practices, the identification of redundant jobs and the improvement of risk management within audited organizations.
- Vanlandingham (2006a), in his doctoral thesis, examined the strategies that were employed by US legislative oversight offices to

stimulate the use of their results in the legislative process. The author distinguished between two overall strategies: developing strong networks with decision makers and astute marketing of products. He particularly examined how key stakeholders utilized and assessed the audits conducted by legislative oversight offices, and to what extent their assessments cohered with the network and marketing initiatives taken by these offices. His research indicated firstly, that offices located within a legislative unit were more effective in promoting use of their work than offices located within an auditing unit. Secondly, he showed that offices that had adopted research standards which stress utility to stakeholders were more successful than offices conforming to the Government Auditing Standards, which strongly emphasize organizational independence (Vanlandingham, 2006b).

- Torgler and Schaltegger (2006) examined how audit courts and local autonomy affect political discussion. Their results indicated that a higher audit court competence and a lower level of centralization correlated with a higher level of political discussion. According to these authors, the results in Switzerland suggest that such institutions help improve citizens' willingness to acquire information and discuss political matters. Furthermore, they found that in the Swiss case, higher audit court competencies had a significantly positive effect on tax morale. They suggest that these help to improve taxpayers' tax morale and thus their intrinsic motivation to pay taxes.

Despite these examples, the assessment of impact is far from straightforward. Some years ago, Lonsdale (1999) studied the impact of performance audits conducted by audit offices. He distinguished different types of impact. Firstly, there may be a quantifiable financial impact in terms of a reduction of expenses 'with real and readily accessible changes in the financing of the area in question' (p. 177). He gives the example of audits carried out by the UK National Audit Office targeting the effects of privatization. Secondly, there may be non-financial impacts, such as improvements in customer satisfaction, reductions in waiting times, or shorter construction periods (p. 178). A third type is qualitative impacts, such as service improvements, while the fourth type relate to 'savings for the citizens'. With regard to the fifth category, that of impact on political topics, Lonsdale observes that 'in general, SAIs tend to stay clear of highly political matters but in a small number of cases performance audit reports have had an impact at the political level [in the 5 countries investigated]'. Overall, however, measuring impacts was considered problematic because, as the author makes clear, 'up to now, our discussion of the

impact of performance audits has been based almost entirely on what SAIs have reported about themselves. Assessing whether these views are reasonable is not easy, as there is little independent information available'.⁵ There are other problems. A number of studies suggest that audits and performance measurement work by no means always generate the effects that were intended, that is, greater effectiveness and efficiency. These studies involved both the private sector (Meyer and O'Shaughnessy, 1993; Meyer and Gupta, 1994), as well as the public sector (for an overview, see Van Thiel and Leeuw, 2002). Other commentators present little evidence to support their claims of influence. For example, Peters (2005) has argued that:

Transformation in governance [which includes accountability mechanisms] has, to some extent, enhanced the capacity of government to perform its tasks but in the process these changes have exacerbated the familiar problems of co-ordination and altered the ways in which accountability can be enforced. There is little doubt that government in 2004 is more efficient in a strictly economic sense than it was in 1984, or 1994, or even the late 1990s. Likewise, many of the changes in government that have been implemented during the reform process have enhanced public participation, especially the involvement of clients of programmes . . . Further, there is some evidence that the public is now more satisfied with the individual services provided by the public sector.

Unfortunately, in Peters' analysis the questions of if, and to what extent, the government has become more 'effective' as a result of the increased attention paid to audits and governance reforms, are neither addressed nor answered.

Funkhouser (2005, and in this book) also states that auditors believe that their work has sparked actual improvements in the performance of government agencies, but has to acknowledge that there has been virtually no scholarly examination of this assertion. He found no more than a few scientific studies, whilst Van Looke and Put in their chapter of this book draw on only 14, some of which stretch back to the 1970s.

Looking at the Dutch, British, and Canadian National Audit Offices, Sterck et al. (2006) studied the impact of their respective performance audits. The presented data show that these audit offices put a lot of work into follow-up studies, justified it seems by the hypothesis that when those audited say they have adopted the recommendations (and are seen to have done so in follow up reports), the desired impact on efficiency and effectiveness has, indeed, taken place. It is not clear whether or not this is true. Put (2005b) has also shown that the factors to which audit offices pay attention in their studies are not, in general, the factors that, according to the prevalent social and behavioural sciences literature (see also Leeuw, 1998), are seen as determinants of efficiency and successful outcomes.

The ambition of this chapter is not to be comprehensive, but to highlight the main lessons from empirical research on the effects of performance audits. The first finding, as has been established earlier, is that the number of empirical studies is small. Secondly, it is not always very clear what is meant by impact. For example, is it:

- the acceptance of the critique and associated recommendations made by the auditors;
- a reduction in the number of administrative or accountability problems discovered;
- changes in governance and the ‘administrative organization’;
- a change in the impact on society of policies that have been proposed;
- something else?

Thirdly, compared to the state of the art in impact or effectiveness evaluations, the studies referred to are methodologically unsophisticated. In order to be able to present valid information on the impact of a policy or other type of policy instrument such as performance auditing, several conditions have to be fulfilled (Rossi et al., 2004; Leeuw and Vaessen, 2010). The most important one is to address (and solve) the attribution problem: can a change in a dependent variable such as an organization’s efficiency or a policy’s effectiveness causally be attributed to the intervention (that is, the performance audit)? As far as we could detect, none of the studies referred to above used a design that addressed this problem, which reduces the validity of statements about the ‘effects’ of performance audits.⁶ A second criterion is that the theory underlying the intervention should be articulated and put to test. The theory (or theories) underlying performance audits are only seldom articulated and put to test (Put, 2005b; Leeuw and Crijns, 2005).

The evidence that performance audits ‘work’ in realizing their goals is therefore limited. Without doubt, there are several indicators which show that audits make a difference, but because the causal relationship between audits and ‘change’ is not made clear, one has to be careful in concluding that audits are producing effects.⁷

THE EXTENT TO WHICH AUDITS HAVE UNINTENDED AND UNDESIRE SIDE-EFFECTS⁸

We now turn to examine what research is available on unintended and undesired side-effects. In the literature, several side-effects of performance

audits are mentioned. Pollitt et al. (1999) indicated that the presence of auditors may considerably enlarge the workload of audited departments.⁹ In addition, their work suggests that some of the auditees, when informed of the possibility of being subject to an audit in the near future, have acted to make hasty introductions of new management practices, in anticipation of the auditor's visit.

Other authors (for example, Leeuw, 2009a) have contended that performance audits may impede innovation and renewal in government by the strong emphasis they put on compliance with procedures ('ossification').¹⁰ An unintended side-effect, linked to this, is the development and excessive use of manuals, guidebooks, protocols and checklists in order to cope with the recommendations of auditors and the procedures they prescribe. It is unclear if such an approach contributes to a more efficient and effective public sector. Short termism is also seen as an unintended side-effect, where auditees become inclined to focus primarily on achieving good results against the performance indicators established for their organization, and lose sight of the long-term policy perspectives (Neely, 1999; Jackson, 2005).

Tunnel vision is also mentioned in the literature. Tunnel vision relates to the situation whereby phenomena that are quantified and measured in the performance measurement system are overstressed at the expense of the unquantified and unmeasured aspects of performance (Smith, 1995; Dolmans and Leeuw, 1997). There are other concepts, like the performance paradox, that refer to the situation where performance auditing (as one form of performance management) does not always result in a more efficient and effective policy, and that one also has to take into account the unintended side-effects of performance measurement.

Trust costs are in part to do with the above-mentioned side-effects, but also with the consequences for the level of trust between auditor and auditee, in particular, when auditors make mistakes, or are not precise enough in their diagnoses. Elliott (2002) suggests that performance auditing may even harm the basis for trust on the work floor, and in that way can be counter-productive for organizations. Bunting's book *Willing Slaves: How the Overwork Culture is Ruling our Lives* (2005) argues a similar case, drawing on Michael Power's Audit Society ideas. She suggests: 'A sense of public esteem used to be an important reward of the job of a public servant, but no longer; the political focus on public sector reform exacerbates the sense of being under scrutiny – and of being found wanting.'

So what conclusions can we draw from this summary? The first one is that, again, empirical information about side-effects of performance audits is limited. We end up largely empty-handed if we look for well-designed

evaluative studies trying to find causal relationships between (types of) performance audits and side-effects. Nevertheless, the fact that a number of side-effects are addressed in the literature, even that which shares the view that performance audits are a priori positive for performance, should alert auditors to the possible risks.

HOW CAN WE EXPLAIN THE OCCURRENCE OF SIDE-EFFECTS?

That brings us to the final question: How come there are unintended side-effects? The first possible explanation is based on the predictability of the factors to which performance auditors pay attention. Examples given by Put (2005b and in this book) and by Leeuw and Crijns (2005) set out the criteria and norms SAIs tend to use. Predictable auditors run the risk of being ‘gamed’,¹¹ because the audited organization will have insights into what is about to happen, on which matters the audit will focus, and can guess some or all of the norms to observe. Accordingly, auditees can adapt their behaviour with the help of registration systems, audit trails, services, applying manpower, and other means. Slowing down the audit process or ‘swamping’ the auditor with material to read are also ‘strategies’. The transaction costs of such behaviour do not contribute to an improvement of the efficiency and effectiveness.

Predictability may also evoke another reaction, which also has a transaction cost attached to it, namely resistance. In essence, this means that staff within organizations being audited resist or dispute the criteria and norms of the auditors. They do so because they think these norms are irrelevant to the improvement of efficiency and effectiveness, either at all, or in their circumstances, or because they have already gained some experience of them and know them to be ineffective. Such a side-effect occurs when, time and again, auditors make the same type of recommendations, based on the same type of knowledge. Again though, it takes time, money and manpower to mobilize resistance, certainly when it has to be done in a bureaucratically acceptable way, and this, in turn, harms the advancement of effectiveness and efficiency.

A second explanation relates to the intellectual depth of audits. Put (2005b) has shown that reports from different national audit offices (in the UK, the Netherlands, and, as far as the number of reports went, to a limited extent, in Belgium) tend to focus and pronounce upon variables relating to organizational structure, prevailing procedures, agreements and (legal) arrangements between divisions, and the achievements that need to be attained. They pay much less attention to ‘soft’ or ‘behavioural’

mechanisms, even though, to a large extent, human and organizational behaviour is guided by exactly such mechanisms as ‘groupthink’, dissonance reduction, the ‘shadow of the future’, and the ‘crowding out’ mechanism. Cognitive biases, such as the fundamental attribution error, are also important factors.¹² In line with Mahoney (2003) and others, like George and Bennett (2004), Mark, Henry and Julnes (2000), and Pawson and Tilley (1997), these mechanisms can be seen as propositions about how underlying entities, processes, or structures operate in particular contexts to generate outcomes of interest. The less performance auditors pay attention to the role that these and other behavioural mechanisms play in realizing the goals that organizations have set (or been set), the more they are vulnerable to criticism of their work.

A third explanation revolves around the methodological quality of the audits. Of course, audits are sometimes reviewed and commented on positively by independent experts, as is the case with the United Kingdom NAO, where universities review draft reports and have commented on methods and other aspects of the work since 1991 (Lonsdale, 2008). Nevertheless, there can be criticisms of the methodological quality. While in the world of evaluation, theory-driven evaluations are becoming more important, theory-driven audits are rarely found. Secondly, while evaluators pursuing knowledge about the effects of interventions struggle with the attribution problem, auditors seem not to address this problem in their own work. Notwithstanding the developments that Lonsdale discusses in Chapter 5 of this book, most audits remain dependent on relatively straightforward means of gathering and analysing evidence.

A fourth explanation relates to the style of some of the audits. Walsh (1996) has characterized the way in which the US General Accounting Office (now the Government Accountability Office) (GAO) once publicized its studies as ‘ambush hearings’; the meetings where the auditors presented their findings ‘took the agency by surprise’. Eventually it was decided that this had to end, in part, because such an approach was unhelpful with an eye to getting the findings accepted and implemented. Landers (1999), who appeared in front of the United Kingdom House of Commons Public Accounts Committee to answer questions on an audit report, complained about the way that he was treated and how ineffective the questioning turned out to be.

Subsequently, a ‘no surprise-approach’ has been introduced, for example, in the UK. There, expectations of the conduct of public sector auditors were set out in a document called ‘What Public Bodies Can Expect from their Auditors’ (Public Audit Forum, 2000), published by the Public Audit Forum. In Canada, something similar has been produced. A key element of these approaches is the attention paid to the issue of

'consultation', although it should be said that there are risks for auditors in taking such an approach and getting too close to auditees. Based on experience with the consultation of evaluands and other stakeholders during audits, evaluations and inspections this author would hesitate to recommend such an approach to auditors. It may lead to what Pollitt (1999) has called 'partnerial' investigations, during which the distinctions between the roles of participant, stakeholder, and assessor are increasingly blurred. This puts at risk the independence of the auditors, which, as Jan-Eric Furubo set out earlier, is a key element of their role.

A fifth and final explanation concerns the problem and the risk of 'ideological capture', where an audit office adopts a particular ideological or value framework apparently without question, and offers reports within this framework. Dahanayake (2008: 18; 23) provided comments in this respect. He seems to suggest that performance auditing as done in Alberta (Canada) is strongly related to the New Public Management philosophy. He refers to Gendron et al. (2001) suggesting that the Audit Office had 'compromised its independence by promoting the government agenda of the NPM framework'. English (2007: 333–4) points to a similar problem in her study of Australian audit reports of PPPs (public private partnerships). 'Perceptions that . . . auditors-general may have compromised their independence through their attempts to foster "good management practice" are reinforced when the performance accountability frameworks they use are devised by them to aid steering mechanism implementation'. Sutherland (2001; 2002) has presented similar evidence in respect of audits in Canada.

WHAT NOW?

Not so long ago, we lived in a world without much accountability and transparency. In such a world, without doubt the development of auditing was important and effective, as has been shown for the Netherlands (Leeuw, 2009). However, at some point in time, the 'Law Of Diminishing Returns' might start to be applicable, although it is unclear what this may mean. Would it mean that there is too much auditing? Does it mean that the relationship between the scale of audit and its effects gets complicated or out of balance? The answers are not easy to give, and part of the explanation may be that auditors are not very good at producing evidence about the impact and side-effects of their work. To increase their effectiveness, the author would like to offer three suggestions. These are:

- (1) *Make performance audits more 'evidence-based'*. There is little sign that performance audits are based on (scientific) evidence in a way

that is linked to what is going on in the field of impact evaluations. There, attention to the (quasi-) experimental method, together with understanding what makes interventions work, has been increasing over recent decades, covering fields such as crime and justice, education, health and developmental aid initiatives. In addition, knowledge repositories summarizing the results of systematic research reviews of experiments are readily available (Hansen and Rieper, 2010). There is scope to learn from this work.

- (2) *Pay more attention to the cognitive-behavioural mechanisms underlying the behaviour of actors in and outside of organizations.* Theory-driven evaluations, in which the often implicit theories underlying policies and programmes that have to be examined are reconstructed and tested, pay attention to these ‘mechanisms’. They have also produced methodologies for the ‘reconstruction’ and ‘testing’ of these ‘theories’. Auditors may profit from following a similar route.
- (3) *Invest in innovative techniques for data collection.* There are serious limitations to surveys, document analysis, and face-to-face interviews. People do not always speak or ‘write’ the truth; people do not always do what they say; people are not always able to express themselves in jargon-free language, and so on. Interesting new data collection approaches should therefore more often appear on the agenda in the world of performance auditors. Making more use of unobtrusive measures, where the people observed are not aware of what is going on, or bringing the ‘research lab’ into everyday life, are two examples. If one wants to check what drivers are actually doing while driving, the methodology of naturalistic driving is another example.¹³ Making use of digital traces people leave can also open up new horizons for auditors and inspectors.

Thirty to forty years ago, hardly anybody talked about performance audits. Now, they take place in many Western countries and increasingly beyond. However, there is no time to relax or sit back.

NOTES

1. Power (1997) coined the expression of an ‘audit explosion’ which he argues took place in the UK and other countries. Martin (2005) reported that ‘between 1997 and 2003 [Labour] created nine major new inspectorates and greatly expanded the reach and remit of at least eight other inspection services. Reflecting this rapid growth in scale and scope, the cost of external inspection of UK public services rose steeply from £250 million in 1997/8 to £550 million in 2002/3’. Willemsen, Leeuw and Leeuw (2008) calculated that in the Netherlands between 1989 and 2005 the program costs of – in

particular – inspectorates tripled. However, since 2007 the central government has cut the personnel and program costs of audit offices by 30% and inspectorates by 20%.

2. A camel's nose is a metaphor for a situation where permitting something small (believed to be undesirable, for example an extra but small new budget claim) will grow bigger and bigger. The late US Senator Barry Goldwater in 1958 used this metaphor as follows: 'This bill and the foregoing remarks of the majority remind me of an old Arabian proverb: "If the camel once gets his nose in the tent, his body will soon follow."' Quoted in Pierce, Patrick Alan; Miller, Donald E. (2004). *Gambling politics; state government and the business of betting*, Lynne Rienner Publishers p. 133.
3. Another way to determine whether the audit work has served its purposes is by the use of follow-up studies carried out by performance auditors, during which they examine whether or not their recommendations have been adopted.
4. It has also been noticed that auditors may deliberately or inadvertently fail to acknowledge the costs of doing what they suggest, in order to play up their recommendations.
5. See for an interesting attempt, the Crerar Review, available at: www.scotland.gov.uk/Resource/Doc/198627/0053093.pdf. Crerar (p. 48) points out that 'the costs of external scrutiny have to be balanced against the benefits, but the benefits and impact of external scrutiny are difficult to assess'.
6. Crerar (2007: 52) is of the opinion that 'there is a need for more rigour in the way costs and benefits are measured, and the need for an improved system of impact assessment'.
7. Schelker (2007) and Blume (2007) have analysed the relationship between characteristics of national audit offices and macro-economic indicators (such as the budget deficit, credit ratings, and level of corruption in a country).
8. This paragraph is partly based on Weets (2008).
9. In the world of accreditation of higher education programs sometimes it is referred to as the behavioural costs of this type of work.
10. Evidence of this side-effect can be found in the world of (higher) education audits and reviews, and is based on surveys of school masters and other educational personnel. However, more often one hears this critique in informal meetings and conversations (when the auditors have left the room).
11. Earlier, Van Thiel and Leeuw (2002) have called this 'learning perversely'.
12. See Astbury & Leeuw (2010), Elster (2007), and Hedstrom & Swedberg (2005) for more information on these and other mechanisms.
13. In *Naturalistic Driving Studies*, researchers observe traffic situations and drivers in a discreet way, by using extremely small cameras and sensors installed in their subjects' vehicles. This observation takes place during daily rides, and without a test leader coming along. The equipment stays in the car for a minimum of several weeks, and sometimes for several months to a year. It turns out that, in this way, participants forget that they are permanently observed, which enables the emergence of a sound picture of 'natural' traffic behaviour, not only under normal circumstances, but also during situations of conflict.

REFERENCES

- Astbury, Brad and Frans L. Leeuw (2010), 'Unpacking black boxes: Mechanisms and theory-building in evaluation', *American Journal of Evaluation*, **31** (3), 363–81.
- Blume, L. and S. Voigt (2007), 'Supreme audit institutions: Supremely superfluous? – A cross country assessment', available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=965434 (accessed 10 April 2010).
- Bouckaert, G. and W. Balk (1999), 'Public productivity measurement: Diseases and cures', *Public Productivity and Management Review*, **15** (2), 229–35.

- Bunting, Madeleine (2005), *Willing Slaves: How the Overwork Culture is Ruling Our Lives*, London: Harper Perennial.
- Clemens, J. et al. (2005), 'Government failure in Canada, 2005 Report. A review of the Auditor General's Reports 1992–2005', Public Policy Sources, A Fraser institute occasional paper, 86.
- Crerar, L. (2007), 'Report of the independent review of regulation, audit, inspection and complaints handling of public services in Scotland', Edinburgh: Scottish Government', available at <http://www.scotland.gov.uk/Resource/Doc/82980/0053065.pdf> (accessed 15 April 2010).
- Dahanayake, S.J. and K. Jacobs (2008), 'New public management (NPM), managerialism, and Value for Money (VFM) audit: A review of empirical literature with special reference to Australia', AFAANZ Conference Paper, available at: <http://www.afaanz.org/openconf/afaanz/papers.php>
- Dolmans, L.J.F. and Leeuw, F.L. (1997), 'Performance auditing', in *Vraagtekens bij de effectiviteit van een onderzoekstraditie*. NIVRA, Naar een doelmatiger overheid. Goed georganiseerd mensenwerk, Den Haag, pp. 81–8.
- Elliott, J. (2002), 'The impact of intensive "value for money" performance auditing in educational systems', *Educational Action Research*, **10** (3), 499–506.
- English, L. (2007), 'Performance audit of Australian public private partnerships: Legitimising government policies or providing independent oversight?', *Financial Accountability and Management*, **23** (3), 313–36.
- Funkhouser, M. (2005), 'Accountability, performance and performance auditing: Reconciling the views of scholars and auditors', Paper to the Accountable Governance: An International Research Colloquium, 20–22 October, Belfast.
- Gendron, Y., D.J. Cooper and B. Townley (2001), 'In the name of accountability: State auditing, independence and new public management', *Accounting, Auditing and Accountability Journal*, **14** (3), 278–310.
- George, A.L. and A. Bennett (2004), *Case Studies and Theory Development in the Social Sciences*, Cambridge, Massachusetts: MIT Press.
- Gray, A. and B. Jenkins (2007), 'Checking out? Accountability and evaluation in the British regulatory state', in Bemelmans-Videc, M.-L., J. Lonsdale and B. Perrin (eds), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Jersey: Transaction Publishers.
- Hansen, Hanne Foss and Olaf Rieper (2010), 'Institutionalization of second-order evidence-producing organizations' in Rieper, Olaf, Frans L. Leeuw and Tom Ling (eds), *The Evidence Book: Concepts, Generation and the Use of Evidence*, New Brunswick, New Jersey: Transaction, pp. 27–52.
- Hedstrom, P. and R. Swedberg (eds) (1998), *Social Mechanisms*, Cambridge: Cambridge University Press.
- Hood, Christopher, Oliver James, B. Guy Peters and Colin Scott (eds) (2004), *Controlling Modern Government: Variety, Commonality and Change*, Cheltenham: Edward Elgar.
- Jackson, A. (2005), 'Falling from a great height: Principles of good practice in performance measurement and the perils of top down determination of performance indicators', *Local Government Studies*, **31** (1), 21–38.
- Johnsen, A., P. Meklin and J. Vakkuri (2001), 'Performance audit in local government: An exploratory study of perceived efficiency of municipal value for money auditing in Finland and Norway', *The European Accounting Review*, **10** (3), 583–99.

- Landers, B. (1999), 'Encounters with the Public Accounts Committee: A personal memoir', *Public Administration*, **77** (1), 195–213.
- Lapsley, I. and C.K.M. Pong (2000), 'Modernization versus problematization: Value-for-money audit in public services', *The European Accounting Review*, **9** (4), 541–67.
- Leeuw, F. (1998), 'Doelmatigheidsonderzoek van de Rekenkamer als regelgeleide organisatielkunde met een rechtssociologisch tintje?', *Recht der Werkelijkheid*, **14**, 35–71.
- Leeuw, F. (2006), 'Over impact en neveneffecten van performance monitoring en auditing', *Vlaams Tijdschrift voor Overheidsmanagement*, **11** (4), 37–44.
- Leeuw, F. (2009), 'Evaluation policy in the Netherlands', *New Directions for Evaluation*, **123**, 87–103.
- Leeuw, F. (2009a), 'Evaluation: a booming business but is it adding value?', *Evaluation Journal of Australasia*, **9** (1), 3–9.
- Leeuw, F. and M. Crijns (2005), 'Normnaleiving door organisaties. Sociologie toegepast op een actueel vraagstuk', *Bestuurskunde*, **14**, 40–49.
- Leeuw, Frans L. and Jos Vaessen (2010), *Impact Evaluations and Development. Nonie Guidance on Impact Evaluation*, Washington DC: World Bank & Nonie.
- Lonsdale, J. (1999), 'Impacts', in Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilka Summa and Marit Waerness, *Performance or Compliance?: Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- Lonsdale, J. (2008), 'Balancing independence and responsiveness', *Evaluation*, **14** (2), 227–48.
- Lonsdale, J. and M.-L. Bemelmans-Videc (2007), 'Accountability – The challenges for two professions', in Bemelmans-Videc, M.-L., J. Lonsdale and B. Perrin (eds), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Brunswick and London: Transaction Publishers.
- Mahoney, J. (2003), 'Tentative answers to questions about causal mechanisms', Paper presented at the annual meeting of the American Political Science Association, Philadelphia, PA, 28 August. Available at http://www.allacademic.com/meta/p62766_index.html
- Mark, M.M., G.T. Henry and G. Julnes (2000), *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*, San Francisco: Jossey-Bass.
- Martin, S. (2005), 'Evaluation, inspection and the improvement agenda: Contrasting fortunes in an era of evidence-based policy-making', *Evaluation*, **11**, 496–504.
- Meyer, M. and K. O'Shaughnessy (1993), 'Organizational design and the performance paradox', in R. Swedberg (ed.), *Explorations in Economic Sociology*, New York: Russell Sage Foundation, pp. 249–78.
- Meyer, M.W. and V. Gupta (1994), 'The performance paradox', *Research in Organizational Behaviour*, **16**, 309–69.
- Milgrom, T. and R. Schwartz (2008), 'Israel's auditor as policy change agent: The case of water policy', *International Journal of Public Administration*, **31** (8), 862–77.
- Neely, A. (1999), 'The performance measurement revolution: Why now and what next?' *International Journal of Operations and Production Management*, **19** (2), 205–28.
- Pawson, Ray and Nick Tilley (1997), *Realistic Evaluation*, London: Sage Publications.
- Peters, B.G. (2005), 'Squaring several circles: Co-ordination, performance and

- accountability', in Drewry, G., C. Greve and T. Tanquerel (eds), *Contracts, Performance Measurement and Accountability in the Public Sector*, Amsterdam: IOS Press.
- Pollitt, C. (1999), 'Stunted by stakeholders? Limits to collaborative evaluation', *Public Policy and Administration*, **14** (2), 77–90.
- Pollitt, C. (2003), 'Performance audit in western Europe: Trends and choices', *Critical Perspectives on Accounting*, **14**, 157–70.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilka Summa and Marit Waerness (1999), *Performance or Compliance? Performance Audit and Public Management in Five Countries*, New York: Oxford University Press.
- Power, Michael (1997), *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press.
- Prosser, M. (2005), Chair, UK Women and Work Commission, EU Gender Equality Conference, Birmingham, 7–8 November.
- Public Audit Forum (2000), 'What public bodies can expect from their auditors', available at www.public-audit-forum.gov.uk/publicat.htm
- Put, V. (2005a), 'Met welke bril kijken auditors naar de werkelijkheid? Normen van rekenhoven bij het beoordelen van de overheid', Die Keuze Uitgeverij, Brussel, 2005.
- Put, V. (2005b), 'Normen in performance audits van rekenkamers'. Een casestudie bij de Algemene Rekenkamer en het National Audit Office, Leuven, K.U. Leuven (Unpublished dissertation).
- RMO (2000), *Aansprekend burgerschap*, Den Haag: SDU.
- Rossi, P.H., M.W. Lipsey and H.E. Freeman (2004), *Evaluation: A Systematic Approach*, Thousand Oaks, CA: Sage Publications.
- Schelker, M. and R. Eichenberger (2007), 'Rethinking public auditing institutions: Empirical evidence from Swiss municipalities', available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1028927 (accessed 17 March 2010).
- Schwartz, R. and John Mayne (2005), *Quality Matters. Seeking Confidence in Evaluating, Auditing, and Performance Reporting*, New Brunswick: Transaction Publishers.
- Smith, P. (1995), 'On the unintended consequences of publishing performance data in the public sector', *International Journal of Public Administration*, **18**, 277–310.
- Sterck, Miekatrien et al. (2006), De praktijk van performance audit: een drielandenstudie, Rapport, Instituut voor de Overheid, Leuven, 2006-08-22.
- Thomas A. Wilson (ed.), *Rationality in Public Policy: Retrospect and Prospect, A Tribute to Douglas G. Hartle*, Canadian Tax Paper No. 104, Toronto: Canadian Tax Foundation, 109–40.
- Sutherland, S.L. (2001), 'Biggest scandal in Canadian history': HRDC audit starts probity war. School of Policy Studies, Working Paper 23, available at: http://www.queensu.ca/sps/researchers/working_papers/ (accessed 17 March 2010).
- Sutherland, S.L. (2002), 'The Office of the Auditor General of Canada. Government in exile?' School of Policy Studies, Working Paper 31, available at: http://www.queensu.ca/sps/researchers/working_papers/
- Torgler, B. and C.A. Schaltegger (2006), 'The determinants of political discussion: How important are audit courts and local autonomy?', *Berkeley Program in Law and Economics, Working Paper Series*, 198, Berkeley, University of California.
- Treasury (2006), 'Treasury Minutes on the Thirteenth to Seventeenth Reports from the Committee of Public Accounts 2005–2006', Seventeenth Report, available at www.official-documents.gov.uk/document/cm67/6743/6743.pdf

- Vaessen, Jos and Frans L. Leeuw (eds) (2010), *Mind the Gap*, New Brunswick and London: Transaction Publishers.
- Van Thiel, Sandra and Frans L. Leeuw (2002), 'The performance paradox in the public sector', *Public Performance and Management Review*, **25**, 267–81.
- Vanlandingham, G.R. (2006a), 'A voice crying in the wilderness – Legislative oversight agencies' efforts to achieve utilization', Florida, The Florida State University, 2006, Unpublished dissertation.
- Vanlandingham, G.R. (2006b), 'A voice crying in the wilderness – Legislative oversight agencies' efforts to achieve utilization', *New Directions for Evaluation*, **112**, 25–39.
- Walsh, A.H. (1996), 'Performance auditing and legislative oversight in the context of public management reform: The US experience', in OECD (ed.), *Performance Auditing and the Modernisation of Government*, Paris: OECD.
- Weets, K. (2008), 'How effective are performance audits? A multiple case study within the local audit office of Rotterdam', Paper presented at the 5th international conference on Accounting, Auditing & Management in Public Sector Reforms (European Institute for Advanced Studies in Management). The Netherlands: Amsterdam. 3–5 September.
- Wilkins, P. and J. Lonsdale (2007), 'Public sector auditing for accountability: New directions, new tricks?', in Bemelmans-Videc, M.-L., J. Lonsdale and B. Perrin (eds), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, Somerset, NJ: Transaction Publishers.
- Wilkinson, J. (2005), 'The growth of performance measurement: Help or hindrance?', in Drewry, G., C. Greve and T. Tanquerel (eds), *Contracts, Performance Measurement and Accountability in the Public Sector*, Amsterdam: IOS Press.
- Willemsen, Frank, Frans L. Leeuw and Bastiaan Leeuw (2008), 'Toezicht en inspectie in maten en soorten', *Tijdschrift voor Criminologie*, **50** (2008), 96–113.
- Williams, Rob (2003), 'When audits fall short: Weak audits will damage voter trust'. Address, Evergreen freedom foundation, 2003. Available at: <http://effwa.org/main/page.php>

12. Impact at local government level: a multiple case study

Katrien Weets

INTRODUCTION

As noted elsewhere in this book, few academic studies have dealt with the impacts of performance audits.¹ Bowerman et al. (2000) stated: ‘In analysing the shape and scale of the “audit society”, it is a striking paradox that the process of audit itself has been audited or evaluated in only limited ways.’ Despite this, many Supreme Audit Institutions (SAIs) have invested heavily in the expansion of their performance audit activities (Morin, 2004). Prompted by this observation, this chapter contains the results of an in-depth analysis of the impacts of three performance audits conducted by the city audit office of Rotterdam (*Rekenkamer Rotterdam*). It complements the more general overviews of earlier chapters, highlighting a methodology for understanding impact better.

The next part briefly situates this research in the existing literature. The third part sets out the framework used to analyse performance audit impacts and discusses the research methodology applied. The fourth part presents the data gathered for each of the selected cases, whilst the fifth part discusses the difficulties associated with examining performance audit impacts. The chapter concludes by summarizing the main research findings.

WHAT DO WE ALREADY KNOW ABOUT THE IMPACT OF PERFORMANCE AUDITS?

As Van Loocke and Put have shown, there have been very few studies of the impacts of performance audits at the central government or local government level.² Moreover, the studies on the indirect impacts of performance auditing (see Dolmans and Leeuw, 1997; Shore and Wright, 1999; Elliott, 2002; Leeuw, 2006) are frequently based on anecdotal evidence.³

Van Loocke and Put’s chapter is based on a systematic literature review of academic studies that examine the impact of performance audits at the

central government level. In their discussion they draw attention to the divergent ways in which the concept of impact is operationalized in the studies subject to their research. The authors distinguish between five types of impact a performance audit may have: instrumental, conceptual, interactive, political-legitimizing and tactical.⁴ They point out that most studies on the impact of performance audits limit themselves to measuring the direct and instrumental use of the information generated. Hence, the authors conclude, a lot of studies disregard the possibility that theoretical perspectives and conceptions can also permeate the policy-making process in indirect and diffuse ways through the influence they exert on the views, ideas, beliefs and attitudes of public managers, decision makers, and so on (see also Weiss, 1979; Weiss and Bucuvalas, 1980; Leviton and Hughes, 1981). All things considered, Van Looke and Put's literature review suggests that performance audits mostly do have an impact, although this impact can manifest itself in many ways: 'slow and subtle' (Morin, 2008), very explicit and quick, or, on the other hand, at times, hardly any impact at all (see, for example, Milgrom and Schwartz, 2008).

Pollitt et al. (1999) recognized that measuring the precise scale of the impact of a performance audit is difficult, if not impossible. These authors did not study impact in a systematic way, but the evidence they gathered suggests that some performance audits do indeed have an impact. After all, in each of the countries examined, performance audits did lead to changes to government activities and, in many cases, substantial savings in public funds. Besides this, the authors drew attention to the possibility that performance audits produce a deterrent effect arising simply from the fact that they exist, as well as a more generalized educational impact on the audited organizations and on the public sector.

According to Pollitt et al., the indicator most commonly used by the five SAIs subject to their research to determine the impact of their performance audit work, was the proportion of recommendations accepted by government. This indicator can be strongly criticized, however, as it only offers a very limited view on performance audit impacts. It provides an indication of the impacts that occur as a consequence of the performance audit results, but does not take into account the impacts that arise from the performance audit process. It conveys a picture of the impacts that occur at the end of the audit cycle, but overlooks the impacts that emerge during the performance audit process and in the long term. Finally, it neglects the possibility that a performance audit can have both intended and unintended consequences (see Kirkhart, 2000, for what she has called 'an integrated theory of influence').

With regard to the factors that influence performance audits' impacts, Van Looke and Put distinguish between three groups: micro-level factors

(factors referring to the characteristics of the audit, such as the relationship between the auditors and the auditees); meso-level factors (factors relating to the characteristics of the audit institution and the audited organizations such as its reputation, and so on); and macro-level factors (factors regarding the public sector within a country such as the institutional context in which audits take place, the auditors' mandate, and so on).

This chapter explores the impacts of three performance audits conducted by auditors of the city audit office of Rotterdam in the Netherlands, and seeks to explain the differences observed between them. Since the macro-level context was identical for all three performance audits, the macro-level factors could be left out of the explanatory model. Similarly, as reporting on the characteristics of the audited organizations would endanger interviewees' anonymity, we chose not to study and report on the meso-level factors relating to these characteristics. In this way, our research design permitted us to focus entirely on the micro-level factors that, according to the literature, affect the impact of performance audits.

RESEARCH APPROACH

In order to analyse the impacts of the selected performance audits on the one hand, and uncover the various factors that affect these impacts, we chose to start from a model advanced by Morin (2001). We took this model as our point of departure for several reasons. Firstly, it takes into account both the impacts of the performance audit results and the impacts of the performance audit process, and it pays attention to the short-term as well as to the long-term impacts of performance audits. It thus permits us to study performance audit impacts in a more comprehensive way than the measures commonly used by SAIs.⁵ Secondly, Morin succeeded in giving her study a sound scientific basis, and providing a structured framework to examine performance audit influences on public administrations.

Morin (2001), like Van der Meer (1999) and de Vries (2000), regards a performance audit as a social influence process, which takes on its full meaning through the relationship arising between the source of influence, the auditor, and the target of influence, the auditee. She conducted a multiple case study, in which she examined closely the influence attempts in six performance audits performed by the Auditor General of Quebec and the Auditor General of Canada. Using qualitative content analysis techniques, Morin identified 14 performance indicators. These show whether or not an intended influence was perceived as less or more successful, and can be grouped into three categories:

Table 12.1 Performance indicators for performance audits

Category	Performance indicator
1. Auditees' perceptions and reactions with regard to auditors' influence attempts	1. Auditees' feelings towards auditors
	2. Sources of auditees' dissatisfaction with regard to performance audit auditors' work
	3. Auditees' reaction to auditors' influence attempt: internalization
	4. Cooperation offered by auditees to auditors (in auditors' opinion)
	5. Auditees' misgivings concerning the legitimacy of auditors' influence attempt
2. Impact on the audited organization	6. Auditees' perception of added value of the performance audit
	7. Auditees' evaluation of auditors' findings
	8. Willingness of auditees to follow up on auditors' recommendations
	9. Evaluation by auditees of auditors' overall performance
	10. Auditees' perception of the usefulness of the performance audit for the audited organization
	11. Changes made by auditees to management practices
	12. Auditees, perceptions of the overall effect of the performance audit
3. Contribution to the public debate	13. Stimulation of debates in parliament
	14. Coverage by the press

Source: Morin (2001).

- (1) the perceptions and reactions of auditees with regard to auditors' influence attempts;
- (2) the impacts generated by the performance audit;
- (3) the contribution of a performance audit to public debate.

The performance indicators advanced by Morin are summarized in Table 12.1.

Furthermore, Morin identified two categories of success factors that explain why influence attempts in performance audits might either succeed or fail. These are:

- (1) factors linked to the performance audit process;
- (2) factors linked to the existence of environmental conditions.

Table 12.2 Success factors in influence attempts

Category	Success factor
1. Factors linked to the performance audit process	1. Auditees' perception of a participative leadership style in auditors
	2. Auditees' perception of a preference for collaboration on the part of auditors
	3. Auditees' perception of power relations between auditors and auditees
	4. Credibility of auditors in eyes of auditees
	5. Auditees' perception of connotation of auditors' modes of influence and type of message
	6. Auditees' degree of influenceability
	7. Auditees' level of commitment
	8. Auditees' level of tolerance to criticism
	9. Degree of fluidity in communications between auditors and auditees
2. Factors linked to the existence of environmental conditions	10. The will at staff level and in the central authority of the organization being audited
	11. Political will
	12. Timing of the performance audit
	13. Major reorganization in the body being audited
	14. Reform at the government level
	15. Place of the activity audited and of the recommendations within the priority scale of the audited organization's management

Source: Morin (2001).

The success factors put forward by Morin are presented in Table 12.2.

Morin concluded that four of the six performance audits she examined were not very successful. As a result, she argued in favour of a more sceptical attitude towards performance audits' impacts than that commonly set out. Subsequently, however, the author has softened her findings on the basis of an analysis of a survey in which she attempted to measure the impact of performance audits, conducted between 1995 and 2002, on the management of organizations in Quebec. Surveys were sent to 99 auditees, who were subjects of the performance audits. The impact of the selected performance audits was studied in ten different ways, drawing upon factors frequently invoked in the literature (for example, value added by performance audits, the relevance of the recommendations made by the auditors). Based on the audited managers' assessments, the author concluded that performance audits did in general exert a positive influence on

the management of audited organizations. Hence, at least partly, performance audits, as practised in Quebec governmental organizations from 1995 to 2002, seemed to have led to a better performance (Morin, 2004).

Inspired by Morin's classification, we have taken 12 of her 14 performance indicators for this research, leaving out only two in our research design: auditees' reaction to auditors' influence attempt; internalization; and evaluation by auditees of auditors' overall performance. Within the limited scope of our research, it was not feasible to develop a sound and valid measurement for internalization. In addition, we assumed that the different performance indicators partly overlap, so that this indicator was already covered to a certain degree by the others, as was the indicator 'evaluation by the auditees of the auditors' overall performance'.

Like Etverk (2002), we added two performance indicators to the twelve: attention gained from other institutions (such as interest groups and universities); and attention gained from the audited organization. We included the first primarily because of its importance in the Dutch context. As Pollitt and Bouckaert (2004) state in their cross-country comparative analysis of new public management reforms: 'Deliberation, consultation and pursuit of compromise and consensus form the deeply rooted basic traits of Dutch political culture.' Additionally, we aimed, as far as is possible, to give readers a complete picture of performance audit impacts and this is also the justification for the second indicator. Table 12.3 sets out the full list of performance indicators.

The success factors we selected for our own research were also mainly based on Morin's work except for auditees' perception of a participating leadership style in auditors; and auditees' perception of connotation of auditors' modes of influence and type of message. Like Etverk (2002) we replaced these indicators by an indicator relating to the role and function adopted by the auditors during the performance audit process.⁶

Finally, independently of Morin's research, we included five variables in our research design, which the wider literature suggests will influence performance audit impact: the auditees' perception of the extent to which a performance audit is relevant to issues the audited organization has to deal with (Leviton and Hughes, 1981); the degree of cognitive coupling between auditors and auditees (De Vries, 2000); auditees' perception of the technical quality of a performance audit; the extent to which a performance audit challenges the status quo (that is, existing assumptions, practice and arrangements) (Weiss and Bucuvalas, 1980); and the presence of potential indirect impacts (ossification, tunnel vision, short-termism, reduction of auditees' motivation and disruption of the daily work scheme of the audited organization) (Dolmans and Leeuw, 1997).⁷ Starting from the variables mentioned in this paragraph, we developed a research

questionnaire. This questionnaire served as the structure for the interviews with the auditors and auditees of the selected performance audits.

Three performance audits, conducted by the city audit office of Rotterdam, were selected for our research and studied in depth. The following criteria were used in order to select the cases. These were:

- The timing of the performance audit. The impacts of a performance audit are ideally examined about three years after it was conducted. This term offers a reasonable assurance that the policy impacts in both the short and the longer term have occurred.
- The subject of the performance audit. An audit ideally deals with a topic of ongoing policy and delivery, if one aspires to examine its impacts. Auditees will have had the opportunity to implement the recommendations of the auditors, which is not always the case if an audit concerns a one-off project.
- Availability of the auditees. From a practical point of view it was important that a sufficient number of auditees still worked for the audited organization. Therefore, the audit should not have been conducted too long ago.

Based on an exploratory interview with an auditor of the city audit office of Rotterdam, three audits were selected that met these selection criteria. To study their impacts, different research techniques were used. Firstly, a documentary analysis was carried out. Secondly, 15 semi-structured interviews were set up with six auditors and nine auditees. The interviews were structured. First, respondents were asked to score the items on a five-point Likert scale. Next, they were invited to justify and/or to elaborate on their answers. The respondents were free to add comments and/or remarks.

Furthermore, the auditees within the administration (such as public sector managers) instead of the auditees within the political system (for example, the mayor and the aldermen) were selected for interview. There were several reasons for this. These were:

- the exploratory interview with the auditor of the city audit office of Rotterdam revealed that, in general, auditors had much more contact with public servants than politicians during the performance audit process. Since Morin considered a performance audit primarily as a social influence process, the choice for interviewing the auditees within the administration resulted logically from her approach; and
- in addition, one has to take into account street-level bureaucracy theory. It is not because the recommendations of auditors are

adopted at the political level that they actually permeate within the 'shop floor practices' of public organizations. Therefore, we assumed that interviewing people closest to the policy implementation process would provide a more complete picture of performance audit impacts than interviewing those with political responsibility for a certain policy area.

RESEARCH FINDINGS

Table 12.3 presents the average scores given by both auditors and auditees to the different performance indicators advanced by Morin (2001) for each performance audit. It shows that, when attributing the same weighting factor to each performance indicator, performance audit 2 was considered the most successful, performance audit 1 came off second best and performance audit 3 was regarded as the least successful.

The high scores that were assigned by the auditees to the performance indicator 'absence of doubts about the legitimacy of auditors' influence attempt' immediately catch the eye. They lead us to conclude that the greater proportion of the auditees considered the influence attempts of the auditors as legitimate. Moreover, none of the auditees indicated they had doubts about the legitimacy of the performance audit. On the contrary, the semi-structured interviews showed that a significant majority of the auditees strongly supported the idea of a city audit office.

The three performance audits also achieved moderately positive to positive scores with regard to the attention gained from audited organizations and the municipal council. A large majority of the auditees declared that the performance audits were considered important by their organizations, especially because of the consequences they could potentially have for the audited divisions. Only one auditee stated that the organization he worked for had paid little attention to the performance audit to which it had been subject. According to this person, there were several reasons for this. Firstly, the conclusions of the audit were not shocking. Secondly, the bureaucratic structures of the organization he worked for did not allow for quick diffusion of the results of the audit. More broadly, a number of auditees also spontaneously referred to the additional workload the performance audits had caused. Furthermore, the three performance audits also gained quite a lot of attention from the municipal council, as they were all discussed both in a council committee and in a plenary session.

Another striking point of similarity between the three selected performance audits is the low score they all got for the attention gained from other institutions. During the interviews, the auditors referred to their efforts to

Table 12.3 Average scores given by auditors and auditees to the performance indicators advanced by Morin (2001) by performance audit

Performance indicator	PA 1	PA 2	PA 3
1. Auditees' feeling towards auditors	3.25	2.50	1.67
2. Auditees' satisfaction with regard to the performance audit auditors' work	3.25	3.00	2.33
3. Cooperation offered by auditees to auditors (in auditors' opinion)	3.00	4.00	3.50
4. Absence of doubts about the legitimacy of auditors' influence attempt	4.25	4.00	4.67
5. Auditees' perception of added value of the performance audit	3.00	4.00	1.67
6. Auditees' evaluation of auditors' findings	4.00	4.00	2.00
7. Willingness of auditees to follow up on auditors' recommendations	2.50	4.25	2.83
8. Auditees' perception of the usefulness of the performance audit	3.00	4.00	2.33
9. Changes made by auditees to management practices	2.00	3.75	2.25
10. Auditees' perceptions of the overall effect of the performance audit	2.63	3.50	3.00
11. Attention from the audited organization	3.50	3.75	3.92
12. Coverage by the press (in auditees' perception)	3.00	3.25	2.92
13. Attention from the municipal council	3.63	4.00	3.83
14. Attention from other institutions	2.38	2.25	2.67
Maximum score	70.00	70.00	70.00
Total score (in absolute numbers)	43.39	50.25	39.59
Total score (in %)	61.99	71.79	56.56

Notes: Average scores on a scale running from 1 to 5, where 5 represented the best result and 1 the poorest. The performance indicators evaluated by the auditees were 1, 2, 4, 5, 6, 8 and 10. Performance indicator 3 was scored by auditors. Both the perceptions of the auditees and the perceptions of the auditors were taken into account for the following performance indicators: 7, 9, 11, 12, 13 and 14.

reach the general public by using several media channels. Yet they had to acknowledge that they did not consciously supplement these efforts with an active communication strategy directed at the immediate stakeholders of the policy area subject to performance audit.

With regard to the other performance indicators, it can be said that their scores differ widely amongst the three cases. Table 12.4 offers a global overview of the results by category of performance indicators.

Table 12.4 Global overview of the results by category of performance indicators

Category of performance indicators	PA 1	PA 2	PA 3
Auditees' perceptions and reactions with regard to auditors' influence attempts	13.75/20.00	13.50/20.00	12.17/20.00
Impacts generated by the performance audit	17.13/30.00	23.50/30.00	14.08/30.00
Contribution to the public debate	12.51/20.00	13.25/20.00	13.34/20.00
Total score	43.39/70.00	50.25/70.00	39.59/70.00

Based on these results, one could easily assume that the higher impact rate of performance audit 2 directly follows from its relatively high score on the second category, 'impacts generated by the performance audit on the audited organization'. This statement should be put into perspective, however. The interviews with the auditees of performance audit 2 revealed that the changes made to the policy processes did not result from the performance audit alone, but also from a wider range of public management reforms that had been going on for a longer period in the municipal administration. Consequently, the auditees viewed the performance audit report as an external source of legitimacy for the new line of policy that they were already following of their own accord.

Furthermore, it is important to mention that the auditees spontaneously defined the concept of impact very broadly. Making use of the questionnaire, a number of perceptions were examined as to whether, in the auditees' view, the performance audit:

- made their organization more cost efficient;
- helped their organization to reach its goals;
- helped it to work in a more result-oriented manner;
- contributed to the computerization within their organization; or
- made policy initiatives more evidence-based.

Although some auditees indicated that the performance audit did have an impact on their organization, they assigned only low scores to the questions dealing with the topics outlined above. During the semi-structured interviews this issue was discussed with auditees, as it seemed rather contradictory at first sight. The greatest number explained their seemingly contradictory behaviour by referring to the positive impact the performance audit had on their awareness of new public management reforms.

Although they acknowledged that these reform processes had been going on for a longer period, they pointed out that the performance audit had, for the first time, really confronted them directly with the implications of the reforms. Table 12.5 presents the scores the auditees attributed to the different success factors put forward by Morin (2001).

It is worth commenting on the relatively low scores the auditees assigned to the success factor 'credibility of the auditors'. The semi-structured interviews revealed that the auditees did not automatically link this success factor to the auditors' professional competences. When assessing auditors'

Table 12.5 Average scores given by auditors and auditees to the success factors

Performance indicator	PA 1	PA 2	PA 3
1. Auditees' perception of a preference for collaboration on the part of auditors	3.00	1.50	3.00
2. Auditees' perception of power relations between auditors and auditees (absence of struggle for power)	2.25	2.50	1.33
3. Credibility of auditors in the eyes of auditees	3.50	2.50	3.67
4. The role and function adopted by the auditors	3.25	1.50	1.00
5. Auditees' degree of 'influenceability'	4.00	3.50	3.50
6. Auditees' level of commitment	4.00	4.00	4.00
7. Auditees' level of tolerance to criticism	2.00	4.00	1.50
8. Degree of fluidity in communications between auditors and auditees	3.25	3.25	3.17
9. The will at staff level and in the central authority of the organization being audited	2.75	4.50	3.00
10. Political will	2.38	4.50	2.42
11. Timing of the performance audit	4.00	4.00	3.17
12. Major reorganization in the body being audited	/	/	3.08
13. Reform at the government level	/	/	3.00
14. Place of the activity audited and of the recommendations within the priority scale of the audited organization's management	3.88	4.75	4.33
Maximum score	60.00	60.00	70.00
Total score (in absolute numbers)	38.26	40.50	40.17
Total score (in %)	63.77	67.50	57.39

Notes: Average scores on a scale running from 1 to 5, where 5 represented the best result and 1 the poorest. The success factors evaluated by the auditees were 1, 2, 3, 4 and 9. Success factors 5, 6 and 7 were scored by auditors. Both the perceptions of the auditees and the perceptions of the auditors were taken into account for the following success factors: 8, 10, 11, 12, 13 and 14.

credibility, on the other hand, they referred to other qualities, such as auditors' empathy with the specific features of the audited organizations. The main criticisms directed against the auditors were that they acted too much according to the letter, rather than the spirit, of the law. In some auditees' opinion, auditors could have presented certain findings in a more positive way if they had taken into account the context in which the audit took place. Finally, some auditees felt that the auditors did not pay enough attention to changes the audited organization had already made on its own initiative.

In addition, the rather low scores for the success factor 'the auditees' perception of power relations between auditors and auditees (absence of struggle for power)' immediately catch the eye. Although they create the impression that a significant proportion of auditees were convinced that there was a struggle for power between themselves and the auditors, this conclusion should be put into perspective. After all, the scores for this success factor not only varied widely between the different cases, but also between the different auditees of the same performance audit. Furthermore, all performance audits received very high scores for the success factor 'auditees' level of commitment'. In order to substantiate their point of view, the auditors referred to the openness and the extent of the cooperation of the auditees, and the fact that they actively discussed with them the issues arising. In one case they specifically referred to the time and energy the audited organizations invested in the audit, and to the conversations between auditors and auditees, which, according to the auditors, had become quite emotional at times.

Other striking points of similarity between the three selected performance audits are the high scores they all gained on the success factors 'timing of the performance audit' and 'place of the activity audited and of the recommendations within the priority scale of the audited organization's management'. With regard to the latter, all the auditees shared the same opinion, namely that the performance audit touched upon the core activities of their organizations. With respect to the former, the majority of auditees viewed the timing of the three performance audits as contributing to their success. The argument most often used by the respondents to justify their statement that the performance audits came at the right moment was the fact that a new policy plan had to be drawn up shortly after the publication of the performance audit report, so that the results of the audit could feed into this new plan. With regard to the other success factors, the scores differ widely along the three cases. Table 12.6 presents these results by category of success factors.

Based on these findings, we can identify for each of the three selected performance audits the factors that fostered their success. Tables 12.5 and

Table 12.6 Global overview of the results by category of success factors

Category of success factors	PA 1	PA 2	PA 3
Factors linked to the performance audit process	25.25/40.00	22.75/40.00	21.17/40.00
Factors linked to the existence of environmental conditions	13.01/20.00	17.75/20.00	19.00/30.00
Total score	38.26/60.00	40.50/60.00	40.17/70.00

12.6 show that these varied strongly across the three performance audits. In general one can state, in accordance with Morin's approach, that the factors linked to the performance audit process and the factors linked to environmental conditions contributed equally to the impact rate of performance audit 1. With regard to performance audits 2 and 3 the environmental factors seemed to play a more important role, in particular in the case of performance audit 2, where their contribution to its success is quite remarkable. Although performance audit 2 did not do well on the factors linked to the performance audit process itself, it emerged as having the highest impact rate thanks to the favourable environmental conditions. The same can be said for performance audit 3, but in this case the factors linked to the existence of environmental conditions only compensated for the moderate scores on the factors linked to the performance audit process to a lesser extent. Finally, as was already outlined, independently of Morin's framework, we also registered the scores for five other variables, which according to the literature affect the impact of a performance audit. Table 12.7 presents the scores auditees and/or auditors attributed to these variables.

Table 12.7 shows clearly that while performance audits 1 and 2 were deemed relevant by the auditees, this was not the case for performance audit 3. The auditees justified their answer by contending that although they considered the audit to be important (for example, because it touched upon their organization's core activities), they did not view the conclusions and recommendations of the auditors as good starting points to improve the functioning of their organization. According to the auditees this was a consequence of the disagreement that existed between auditors and auditees about the way in which an audit should be conducted, as well as the way in which policy-making and policy implementation should be organized. Auditees particularly criticized the fact that the auditors focused on the letter of the law, instead of its spirit. Moreover, they felt that the auditors laid too much emphasis on quantifying goals, which in their view was simply not possible in the policy areas in which their organization was active. As Table 12.7 demonstrates, these areas of conflict between the auditees and

Table 12.7 Average scores for other variables put forward in the literature as exerting an influence on the impact of a performance audit

Variable	Average score PA	Average score PA	Average score PA
	(in %) 1	(in %) 2	(in %) 3
1. Auditees' perception of the relevance of a performance audit to issues the audited organization has to deal with	70.00	80.00	43.40
2. Degree of cognitive coupling between auditors and auditees	60.00	55.00	20.00
3. Auditees' perception of the technical quality of a performance audit	85.00	67.50	63.33
4. The extent to which a performance audit challenges the status quo (i.e. existing assumptions, practice and arrangements)	54.38	60.00	41.67
5. Presence of potential side effects (ossification, tunnel vision, short-termism, reduction of auditees' motivation and disruption of the daily work scheme of the audited organization) according to the auditees	59.00	54.00	70.67

auditors of performance audit 3 also manifested themselves in their scorings of the degree of cognitive coupling between auditors and auditees.

The technical quality, on the other hand, was quite positively assessed by the auditees of all three selected performance audits. However, it is notable that, while performance audit 2 came out as having the highest impact rate according to Morin's approach, it did not get the highest score on this particular feature. This could be an indication that the technical quality of a performance audit is not an important determinant of its impact. This view was confirmed by the statements of the auditees during the interviews. An example can illustrate this. Although the auditees of performance audit 2 were very positive about the conclusions of the auditors, they were not completely satisfied with the way in which the audit was conducted. In addition, they indicated that the performance audit only partly covered the policy area in which their organization was active, so that it could not grasp fully its complexity. The auditees of performance audit 3 for their part acknowledged that the auditors were very competent and very professional. In their opinion, however, the auditors' lack of empathy harmed the technical quality of the audit and its conclusions.

Furthermore, Table 12.7 shows that the three performance audits are also quite similar in their scores for the extent to which they were seen as a challenge to the status quo by the auditees. The majority of the auditees of performance audits 1 and 2 reported that these audits had had an effect on the existing conceptions and assumptions within their organization. In order to substantiate these statements, they mainly referred to the New Public Management body of thought (for example, the idea of a more competitive, results-oriented and customer-driven government) that as a consequence of the audit would have been incorporated more solidly within their organizations.

Finally, for all three performance audits we verified to what extent certain potential indirect impacts (ossification, tunnel vision, short-termism, reduction of auditees' motivation, and disruption of the daily work scheme of the audited organization), already discussed by Leeuw in Chapter 11, manifested themselves in the auditees' perception. The results in Table 12.7 indicate that indirect impacts were especially present in the case of performance audit 3. According to the auditees of this performance audit, three indirect impacts appeared: disruption of the daily work scheme of the audited organization, ossification and tunnel vision. Particularly remarkable was their answer to our question regarding short-termism. The auditees of performance audit 3 stated that as a result of the performance audit, they actually felt more inclined to pursue their long-term goals, and that in fact the opposite of short-termism had occurred. They contended more specifically that, as a consequence of the negative criticisms of the auditors, they had had to defend themselves and their organization, which had encouraged them to more clearly formulate the organization's mission and its long-term goals.

DIFFICULTIES RELATED TO EXAMINING PERFORMANCE AUDIT IMPACTS AND LIMITATIONS TO THE RESEARCH PRESENTED HERE

The results of our study seem to corroborate Morin's (2001) findings. Of course, we should put this into perspective since we only examined three cases. In the course of conducting this research, we encountered some limitations to our research approach in particular, and to research dealing with performance audit impacts in general.

Firstly, readers should be aware that by using the research approach applied above, one does not measure the impact of performance audit in terms of better policy and better public management, but rather in

terms of the success of auditors' influence attempts. It should be clear that our aim was to examine the impacts of performance audits, and not the impacts of the policy measures subject to these performance audits. As such, by using this method, one cannot guarantee that the implementation of the auditors' recommendations truly contributed to more efficient, effective and transparent policy and management processes.

Secondly, it should be noted that, using this method, one cannot state for certain that, when auditors had recommended certain alterations, and the audited organization had implemented these changes, the audited organization had acted solely as a result of the performance audit, rather than as a result of other events. During the semi-structured interviews, some auditees, for example, indicated that organizational reforms had been carried out as a consequence of new public-management thinking that was permeating the municipal administration, rather than as an immediate result of one of the performance audits. It is important to take into account such methodological issues in the future if one wants to make progress in studying and measuring performance audit impacts.

Thirdly, the research method applied above does not rule out the possibility that rather than a single performance audit having an effect, there could be a cumulative effect of several performance audits. This possibility could not be considered in our research, since the interviewed auditees did not have any previous experiences of the city audit office of Rotterdam. However, future research should take into account the possibility of performance audits having cumulative effects, in order to secure a comprehensive picture of performance audit impacts.

Fourthly, one should be careful in dealing with causality. Often the literature on performance audit impact assumes linear causal relationships as a given. Consequently, other possible causal relationships are often already precluded in advance. However, as the following examples demonstrate, by using such a unilinear approach towards causality, one risks misrepresenting the extent to which performance audits truly have an impact. By using the indicators 'the amount of attention a performance audit gained in parliament (here: municipal council)' and 'the amount of press coverage a performance audit received' one, for instance, implicitly assumes that the more impact a performance audit has, the more attention will be gained in parliament and the more press coverage will be secured. However, this reasoning can equally well be reversed. One could just as well state that the more attention a performance audit gains in parliament and the more press coverage a performance audit gets, the more impact it will have.

In addition, circular causality could also arise. Circular causality occurs when a factor A co-determines the condition of a factor B, which for its part co-determines the condition of factor A. We can illustrate this somewhat

abstract statement with an example. If the media and the parliament, for instance based on their previous experiences, expect that a performance audit will have a big impact, it is plausible that they will devote more attention to this particular audit than if they did not have these expectations. The attention gained from the press and from the media could then foster the performance audit's success, and, as such, enlarge its impact.

Moreover, the possibility that third-party factors influence the relations between certain independent variables and the dependent variable, in this context performance audit impacts, cannot be precluded. It is, for instance, possible that media attention and/or attention from the parliament in itself did not affect performance audit impact, but that these factors stimulated the audited organization to follow up a performance audit closely, which in the end led to a more effective audit. The attention a performance audit gained from parliament and/or media should then be seen as an intermediary variable. It may not affect performance audit impact in a direct way, but rather exert influence in an indirect manner.

Finally, most studies on the impact of performance audits, including this one, do not take into account the preventive effects performance audits produce according to some authors (see for example, Pollitt et al., 1999; Morin, 2004). In their view, performance audits have a deterrent effect on audited managers, in that they discourage them from setting up hazardous management activities. If we eventually want to be able to fully grasp the extent to which performance audits have an impact, more research will be needed on this topic.

CONCLUSION

Our exploratory research had three main ambitions. Firstly, it sought to explore the extent to which performance audits truly have an impact. Secondly, it aimed at uncovering the various factors that affect the extent to which a performance audit has an impact. Thirdly, and no less important, it was designed to re-open the debate on performance audit impacts on the one hand, and the way in which they should be studied on the other.

We argued against the 'popular' indicators that are nowadays most commonly used by SAIs to determine the impact of their performance audit work, as they only measure performance audit impact in a limited way. Inspired by Morin's (2001) approach to performance audits' success, we selected three performance audits, conducted by auditors of the city audit office of Rotterdam, and we applied her more comprehensive model to them. Our research results appear to corroborate Morin's (2001) findings. Based on our findings and our experiences, we brought to the fore a

number of limitations of the research method we used in particular and of research dealing with performance audit impacts in general.

NOTES

1. In accordance with the current practice in the field of policy evaluation, an 'impact' is defined as the result of a certain policy program, intervention or measure (within this particular context: a performance audit). A 'specific' impact refers to the intended results in regard to the direct target groups. An 'indirect impact' relates to the consequences that appear but which were not originally intended.
2. But see Brown, 1980; Hatherly and Parker, 1988; Johnston, 1988; Van der Meer, 1999; Morin, 2001 and 2004; de Vries, 2000; Lapsley and Pong, 2000; Alon, 2007; Schelker, 2008; Milgrom and Schwartz, 2008 at central government level, and Johnsen et al., 2001; Torgler, 2005; Torgler and Schaltegger, 2006; Schelker and Eichenberger, 2008 at local government level.
3. See Weets (2008) for a more extensive discussion of the literature on this topic.
4. A performance audit can be said to have an 'instrumental impact' when the information it brings forward is used in a direct and assignable way (e.g. the number of recommendations adopted by an audited organization). 'Conceptual impact' refers to the situation where a performance audit influences the policy-making process in an indirect manner, by exerting influence on the views, the frames of reference, and the ideas of politicians and managers. 'Interactive impact' relates to the influence a performance audit can have on the relations between the administration, the Cabinet of a Minister, and interest groups as well as to the impact it may have on other 'knowledge creators' (e.g. think tanks, universities, etc.). The 'political-legitimizing impact' of a performance audit can be said to refer to the attention a performance audit gains in parliament and/or media. The 'tactical impact' of a performance audit can manifest itself in different ways. The results of a performance audit may for example be (mis)used during the consultations between departments and agencies. These different types of impact are not mutually exhaustive. Hence, they do not exclude each other.
5. See Weets (2008) for a more detailed review of arguments against the 'popular' indicators SAIs generally use to measure the effectiveness of their performance audit work.
6. See Pollitt et al. (1999) for an elaborate discussion of performance auditors' roles.
7. Tunnel vision relates to the situation whereby phenomena that are quantified and measured in the performance measurement system are overstressed at the expense of the unquantified and unmeasured aspects of performance (Smith, 1995; Dolmans and Leeuw, 1997). Ossification refers to the situation whereby performance audits impede innovation and renewal in government by the strong emphasis they put on compliance with procedures (Leeuw, 2006). Short-termism relates to the situation whereby auditees as a result of performance auditing become inclined to focus primarily on achieving good results with regard to the performance indicators used by their organization and lose sight of the long-term policy perspectives (Neely, 1999; Jackson, 2005).

REFERENCES

- Alon, G. (2007), 'State audit and the media as the watchdogs of democracy – a comparative view', *Journal of the Office of the State Comptroller and Ombudsman of Israel*, **61**, 55–100.
- Bowerman, M., H. Raby and C. Humphrey (2000), 'In search of the audit society: some evidence from health care, police and schools', *International Journal of Auditing*, **4**, 71–100.

- Brown, R. (1980), 'Auditing and public administration: the unrealized partnership', *Public Administration Review*, **40** (3), 259–65.
- Dolmans, L.J.F. and F.L. Leeuw (1997), 'Performance auditing. Vraagtekens bij de effectiviteit van een onderzoekstraditie', *NIVRA, Naar een doelmatiger overheid. Goed georganiseerd mensenwerk*, 81–8.
- Elliott, J. (2002), 'The impact of intensive "value for money" performance auditing in educational systems', *Educational Action Research*, **10** (3), 499–506.
- Etverk, J. (2002), 'Measuring performance audit effectiveness: the case of Estonia', Master Thesis, University of Tartu.
- Hatherly, D.J. and L.D. Parker (1988), 'Performance auditing outcomes: a comparative study', *Financial Accountability and Management*, **4** (1), 21–41.
- Jackson, A. (2005), 'Falling from a great height: principles of good practice in performance measurement and the perils of top down determination of performance indicators', *Local Government Studies*, **31** (1), 21–38.
- Johnsen, A., P. Meklin and J. Vakkuri (2001), 'Performance audit in local government: an exploratory study of perceived efficiency of municipal value for money auditing in Finland and Norway', *The European Accounting Review*, **10** (3), 583–99.
- Johnston, W.P. (1988), 'Increasing evaluation use: some observations based on the results at the U.S. GAO', *New Directions for Program Evaluation*, **39**, 75–84.
- Kirkhart, K.E. (2000), 'Reconceptualizing evaluation use: an integrated theory of influence', *New Directions for Evaluation*, **88**, 5–23.
- Lapsley, I. and C.K.M. Pong (2000), 'Modernization versus problematization: value-for-money audit in public services', *The European Accounting Review*, **9** (4), 541–67.
- Leeuw, F.L. (2006), 'Over impact en neveneffecten van performance monitoring en auditing', *Vlaams Tijdschrift voor Overheidsmanagement*, **11** (4), 37–44.
- Leviton, L.C. and E.F.X. Hughes (1981), 'Research on the utilization of evaluations – a review and synthesis', *Evaluation Review*, **5** (4), 525–48.
- Milgrom, T. and R. Schwartz (2008), 'Israel's auditor as policy change agent: the case of water policy', *International Journal of Public Administration*, **31** (8), 862–77.
- Morin, D. (2001), 'Influence of value for money audit on public administrations: looking beyond appearances', *Financial Accountability and Management*, **17** (2), 99–118.
- Morin, D. (2004), 'Measuring the impact of value-for-money audits: a model for surveying audited managers', *Canadian Public Administration*, **47** (2), 141–64.
- Morin, D. (2008), 'Auditors generals' universe revisited. An exploratory study of the influence they exert on public administration through their value for money audits', *Managerial Auditing Journal*, **23** (7), 697–720.
- Neely, A. (1999), 'The performance measurement revolution: why now and what next?', *International Journal of Operations and Production Management*, **19** (2), 205–28.
- Pollitt, Christopher and Gert Bouckaert (2004), *Public Management Reform: A Comparative Analysis*, Oxford: Oxford University Press.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilka Summa and Marit Waerness (1999), *Performance or Compliance? Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- Schelker, M. (2008), 'Public auditors: empirical evidence from the US States', CREMA Working Papers 2008/04, Basel: Center for Research in Economics, Management and the Arts.

- Schelker, M. and R. Eichenberger (2008), 'Rethinking public auditing institutions: empirical evidence from Swiss municipalities', CREMA Working Papers 2008/06, Basel: Center for Research in Economics, Management and the Arts.
- Shore, C. and S. Wright (1999), 'Audit culture and anthropology: Neo-liberalism in British higher education', *Journal of the Royal Anthropological Institute*, **5** (4), 557–75.
- Smith, P. (1995), 'On the unintended consequences of publishing performance data in the public sector', *International Journal of Public Administration*, **18** (2), 277–305.
- Togler, B. (2005), 'A knight without a sword or a toothless tiger? The effects of audit courts on tax morale in Switzerland', Working Paper No. 2004/06. Center for Research in Economics, Management and the Arts.
- Togler, B. and C.A. Schaltegger (2006), 'The determinants of political discussion: how important are audit courts and local autonomy?', in *Berkeley Program in Law and Economics, Working Paper Series*, 198, Berkeley: University of California.
- Van der Meer, F.B. (1999), 'Evaluation and the social construction of impacts', *Evaluation*, **5** (4), 387–406.
- Vries, G.J. de (2000), 'Beleidsdynamica als sociale constructie. Een onderzoek naar de doorwerking van beleidsevaluatie en beleidsadvisering', Delft: Uitgeverij Eburon.
- Weets, K. (2008), 'How effective are performance audits? A multiple case study within the local audit office of Rotterdam', Paper presented at the 5th international conference on Accounting, Auditing and Management in Public Sector Reforms (European Institute for Advanced Studies in Management), The Netherlands: Amsterdam. 3–5 September 2008.
- Weiss, C.H. (1979), 'The many meanings of research utilization', *Public Administration Review*, **39** (5), 426–31.
- Weiss, Carol and Michael Bucuvalas (1980), *Social Science Research and Decision-Making*, New York: Columbia University Press.

13. Learning in an accountability setting

Jeremy Lonsdale and Elena Bechberger

This book has emphasised that the primary purpose of performance audit is to help secure accountability, but has also highlighted the value of learning from the work. To some (Lehtonen, 2005), reconciling the demands of accountability and learning can be problematic, and others (Bemelmans-Videc, Perrin and Lonsdale, 2007) argue that ‘traditional forms of accountability.[are] . . . often viewed as less concerned with learning than with punishment’. Some have noted that a primary focus on accountability brings with it a strong focus on rigour, independence, replicability and efficiency, whereas a focus on learning emphasises stakeholder ‘buy-in’ and an evaluation process which leaves space for discussion and lesson-drawing. It has been concluded ‘These two objectives are not necessarily incompatible . . . but they are sufficiently different to merit separate consideration’ (OECD, 2001).

Generally, performance audit is associated with accountability practices, with a normative perspective, and often with attention to weakness and shortcomings. However, another perspective has been heard, primarily from practitioners (Funkhouser in this volume; Black, 2000; Bourn, 2007), emphasising the role they believe audit plays in improving government performance and helping organisations learn. Others have drawn attention to what they see as growing interest amongst auditors in identifying and promulgating ‘good practice’ (Wilkins and Lonsdale, 2007, Bechberger and Page, 2007). Against this background, this chapter examines how performance auditors are seeking to reconcile the demands of helping government to learn with their role in an accountability setting. In particular, it focuses on the National Audit Office (NAO) in the United Kingdom. In the absence of previous considerations of performance auditing and learning, it starts by outlining what the wider evaluation literature considers are the factors which make for effective learning. It then goes on to analyse to what extent the work of the NAO meets those characteristics and considers the resulting challenges of seeking to combine accountability with learning in performance audit.

Table 13.1 Some factors influencing the success of evaluation in assisting learning

Some factors likely to assist learning	Some factors likely to inhibit learning
<ul style="list-style-type: none"> ● Suitable timing of findings ● Desire to learn ● Capacity to learn ● Established culture of evaluation and monitoring in place ● Evaluation seen as high quality ● On-going links between evaluator and those evaluated 	<ul style="list-style-type: none"> ● Evaluation seen as a threat ● Conclusions and recommendations seen as unrealistic ● Ideological opposition ● Poor dissemination of findings ● Concern at potential embarrassment, risk to reputation, etc. ● Lack of trust of evaluator ● Prior bad experiences of evaluation

EVALUATIVE ACTIVITY AND LEARNING

A learning government has been described as one ‘capable of improving its policy measures and underlying assumptions or policy theories’ (van der Knaap, 2006). This can happen, for example, through doing and then reflecting on the experience, through staff attending courses or sharing experiences with peers, through having approaches demonstrated or explained; through challenge, constructive or otherwise; through comparisons with other organisations or between units; and through external scrutiny, audit, inspection or evaluation. There is widespread discussion in the literature of how this evaluative work may contribute to learning in government. Howlett and Ramesh (1995), for example, see policy evaluation as ‘part of a process of learning in which policies develop and change largely on the basis of conscious recognition of past successes and failures and conscious efforts to emulate successes and avoid failure.’

Successful learning from evaluation is, of course, far from inevitable or straightforward, and may be dependent on a range of factors including organisational capacity, the approaches used, the authority of those carrying out evaluations, the appropriateness of timing, luck and whether there are forces working against learning (Table 13.1).

Some commentators (see, for example, Preskill and Torres, 2000) have emphasised the importance of a collaborative approach (‘participatory, dialogic, reflective and inquiry-oriented approach to evaluation and the use of findings’) as the best way to lead to learning. These authors have highlighted the importance of ongoing links between evaluators and those evaluated; the incremental and iterative nature of learning; and the value of learning from past evaluations.

On the other hand, there can be barriers. Taut and Brauns (2003) propose a framework for thinking about stakeholder resistance to programme evaluation, which we believe is also relevant to performance audit. They argue that a basic need for control can lead to a defence of the status quo. The need for a positive self-image can lead to the rejection of performance feedback and to engagement in unco-operative behaviour. Prior experiences also fundamentally influence attitudes towards evaluation, and there is a basic human tendency to maximise rewards while avoiding punishment. In addition, as Rossi et al. (1999) have suggested, 'No matter how well an evaluator's conclusions about the effectiveness of a program are grounded in rigorous research design and sensitively analyzed data, some stakeholders are likely to perceive the results of an evaluation to be arbitrary or capricious judgements and to react accordingly.'

LEARNING IN AN ACCOUNTABILITY SETTING

This general discussion has identified circumstances where learning from evaluative work is more likely. However, such conducive circumstances may not exist or may be very hard to secure and, instead, a series of barriers to learning may be in place. In the worst cases, there may be a track record of difficult interventions, suspicion about motives and defensiveness about the timing or scope of the work. One of those settings may be where there is a strong accountability element to the work (Bemelmans-Videc, Lonsdale and Perrin, 2007). By this we mean where reports and evaluative information are used for formal purposes of holding to account, such as reporting back to grant funders or – as is often the case for performance audits – to parliamentary committees.

For some the development of performance measurement and reporting regimes and of audit are features of a culture that creates problems for learning. It has been argued that:

Such a culture does not lead to the development of learning organizations, able to respond to the needs of customers and citizens, but the establishment of a monitoring and audit culture. It is a stunted, withered thing in which officials are afraid to make innovative decisions and managers enforce obeisance to the latest centrally directed performance targets. (Massey and Pyper, 2005)

Evaluations in accountability settings are considered to have particular effects on officials and induce certain behaviours. For example, some believe the public nature of such work can create problems.

Evaluations that primarily have an accountability function tend to be public and are often performed by external evaluators. Units whose policies, management or implementation activities are being evaluated will often be inclined to defend their actions and achievements. This limits opportunities to learn from evaluation in terms of possible improvements. (van der Meer and Edelenbos, 2006)

This has been contrasted with evaluations outside the realm of accountability where:

Case studies and interview transcripts can be anonymized so that ministers, officials and local managers can own up to, and learn from, mistakes. Skilful evaluators can act as 'honest brokers' and may be able to provide 'safe' environments in which policy-makers and practitioners can suspend the posturing fostered by inspection in order to share insights 'off the record'. (Martin, 2005)

Not everyone agrees that scrutiny for accountability purposes is incompatible with learning. Weiss (1998) has suggested 'When evaluation results are reported to policy makers, funders, and the public, they get a wider hearing. By virtue of being public, results may stimulate adaptive action. Knowing that other people will be looking may make practitioners more attentive to the public's expectations.' Bovens (2004) has also highlighted that accountability is not just about control but also about prevention and, for example, parliamentary inquiries into policy failures can lead 'large numbers of policy managers in similar positions to adjust their policies and procedures.' And it seems, there are occasions when perceptions of audit's role in relation to learning can also change. Gendron, Cooper and Townley (2007) highlight how the Office of the Auditor General in Alberta, Canada became recognised as expert in performance measurement, in the face of scepticism and opposition. The office did this by building up expertise and presenting its views on how performance measurement could be implemented.

Despite these more positive comments, we would argue that, in general, accountability settings have fewer of the characteristics often associated with learning, and more of those which are seen as barriers. This is a challenge for performance auditors seeking to generate learning from their work as it takes place predominantly in an accountability setting.

THE NATIONAL AUDIT OFFICE AND LEARNING

The next section addresses the questions raised by our analysis so far by considering them in the context of the specific case of the National Audit Office (NAO) in the United Kingdom. It examines:

- (1) the existing evidence we have on the NAO and learning;
- (2) what the NAO has said officially about its role with regard to helping government learn;
- (3) the views of experienced performance audit staff on opportunities to help learning; and
- (4) the promotion of 'good practice' by auditors.

What Evidence Do We Have Already?

Many aspects of the circumstances in which the NAO works would seem to favour an effective contribution to government learning. For example, it has a cross-government remit, which allows it to gain a view on the whole of government and make comparisons. It is well-resourced and employs staff from a range of disciplines. It has the ability to report publicly, including synthesising the messages from its own reports. In addition, as mentioned in Chapter 1, its work feeds directly into arrangements by which government bodies are required to explain what actions they are taking to improve performance. The NAO's reports are the basis for hearings of the Committee of Public Accounts, which has strong powers of inquiry and the ability to secure actions against its recommendations.

The NAO also appears to believe it should assist government to learn. Over a decade ago Pollitt et al. (1999) highlighted the importance of the NAO's aspiration, albeit secondary to its accountability responsibilities, of helping public bodies achieve value for money through playing more of a 'consultancy' and advisory role. More recently, Wilkins and Lonsdale (2007) have highlighted the widening of the range of work undertaken by audit offices, including the NAO, in terms of disseminating good practice to government bodies. This is confirmed by the NAO's public statement that 'Our audit of government not only holds its departments and bodies to account, but also helps and supports them in pursuing improvement. That has long been our core conviction and it rests on solid evidence (National Audit Office, 2010).

On the other hand, some of the factors associated with learning identified earlier in the chapter do not appear to be present, and arguably many of the factors deemed to obstruct learning are. For example, the NAO has a reputation for commenting critically on shortcomings and is seen by many officials as having a negative perspective. As an audit body it is strongly associated with what Furubo in his chapter referred to as the 'bad guys' perspective. The NAO is part of formal accountability arrangements that have been described by some as aggressive and critical (Landers, 1999, Lipsey, 2000), and where attendance at a committee hearing becomes a 'damage-limitation exercise for the audited body' (Bowerman, 1996). It

is also outside of government rather than within it (deliberately so, to set it up away from the executive), which may reduce the scope for the kind of informal and on-going collaboration considered important to learning from evaluation. Such distance – whilst deemed important for safeguarding independence – can be interpreted by officials as unhelpful and shape their interpretation of the office as an unlikely source of assistance.

The NAO also has a very wide remit, which arguably limits its ability to return too often to the same subject. As a result, some have questioned whether there is sufficient evidence that, in response to NAO reports, ‘long term improvements in outcomes are achieved, wider lessons are learned and mistakes not repeated’ (Brazier and Ram, 2006). In addition, the importance placed on the independence of the auditor may reduce the willingness or ability of audit staff to play a role in implementation of recommendations, and reduces the chances of an on-going relationship with officials within departments. Its remit does not allow it to comment on the merits of policy, which may be seen as placing an artificial constraint on its ability to assist with learning.

Finally, perceptions of the NAO may affect whether it is seen as an organisation concerned with learning. Pollitt et al. (1999) identified four possible ‘roles’ for performance auditors, including ‘public accountant’ and ‘management consultant’. They suggest that individual SAIs tend not to fit neatly into one ‘role’. Sharma (2007) has highlighted what she regards as:

the differences in both talk and texts between the discourse VFM auditors adopt when interacting with the auditees and that used in their interactions with PAC . . . On the one hand, they act in ‘consultative roles’ and maintain good working relationships with the departments (Bowerman, 1996, Bowerman et al, 2003; and Gendron et al 2001) and, on the other, they manage impressions with the PAC by providing them with critical briefings upon which to question the departments in the hearing.

In these circumstances officials could be excused for being uncertain as to which mode the auditor is operating in at any particular time, or for being concerned that although the auditors may be interested in assisting improvement, their parliamentary allies are more concerned with holding them to account. Table 13.2 examines the nature of the NAO’s approach to performance audit against the factors shaping learning discussed at the start of the chapter.

What Does the NAO Say about Learning and Has this Changed over Time?

We now turn to what the NAO says about its role in relation to learning, based on an examination of NAO Annual Reports from 1988 to 2010 to

Table 13.2 How the National Audit Office matches against factors shaping learning

Factors shaping learning from evaluation	NAO performance audit
Involvement of stakeholders in the evaluation process – collaborative and guided by dialogue and reflection	Traditionally, maintained a distance and, at most, consulted about methods, focus etc. Dialogue often undertaken in formal process of clearing report
Continuous involvement with evaluation audiences to identify the ‘teachable moments’ as well as plan for learning	Typically, undertake work and then move on to another topic or organisation. Some on-going relations maintained where study part of a programme or where follow-up work being undertaken
Iterative nature of transformative learning to ensure that aware of what aspects of the process facilitated or prevented learning	Audit process traditionally focused on specific outputs – opinions, reports – with clear end to involvement
Continuously returning to learning group and encouraging learning as routine part of getting work done	Audit tends to be set piece with traditionally limited involvement post-reporting, except for follow-up work
Blurring of boundaries as evaluators guide program staff in dialogue and reflection processes	Generally strict rules adhered to regarding involvement in implementation. Clear demarcation between auditors and auditees recognised on all sides
The longer the evaluator maintains a relationship with potential users, the greater the likelihood that long-term use will be observed	Long-term relationships do exist and are incentives in performance targets to follow-up and demonstrate impact

identify the scale and nature of the discussion about learning (Table 13.3). To do this, the text of the reports was reviewed to identify the occasions where it referred to going beyond simply executing its statutory duties to report on audit findings. This included the use of the word ‘learning’, ‘lessons’, or ‘disseminating good practice’. Specific outputs or initiatives were also noted.

Table 13.3 gives some indication – albeit drawing on documents which are inevitably seeking to present a favourable view – of what the NAO has reported it has done in order to contribute to learning. These are all examples of what Rossi, Freeman and Lipsey (1999) call ‘secondary

Table 13.3 References to learning activities in NAO Annual Reports 1988–2010

Year	References to lessons and learning activities	Learning activities mentioned
1988	0	
1989	2	Good practice guides
1990	6	Seminars and conferences Presentations and articles
1991	5	Separate reports to supplement published reports Centre of expertise on HR
1992	2	Reports to audited bodies on back of studies Articles, lectures and seminars
1993	4	Articles, lectures and seminars Unit to guide quality of service aspects of VFM work
1994	3	Articles, lectures and seminars. Presentations Treasury letter drawing attention to lessons from NAO
1995	3	Articles, lectures, seminars and visits. Presentations
1996	10	Advise on specific projects Reports for audited bodies Seminars, lectures and other events. Guidance for school governing bodies. Reports to individual schools. Guides with funding bodies. Secondments
1997	7	Seminars, lectures and other events. NAO Focus magazine
1998	3	NAO Focus, checklists from studies. Guide for new defence agencies
1999	5	Leaflet advising people how to save money on their gas bills. Checklists from studies. NAO Focus
2000	6	NAO Focus, seminars
2001	9	Leaflet on domestic electricity consumption. Questions for patients to ask prior to hip operations. Seminars, conferences, speeches. Focus. On-line Focus. Chairing advisory group in a hospital
2002	7	Conferences, seminars. NAO Focus, book on HAI. Checklists in studies. On-line Focus. Included two page section on good practice
2003	6	Conferences, seminars. NAO Focus, on-line, summarise key points for NHS senior staff

Table 13.3 (continued)

Year	References to lessons and learning activities	Learning activities mentioned
2004	10	Support to audit committees, forums for members, best practice guide. Recommendations database for PFI. Training workshops for PFI teams. Conferences, briefings
2005	20	Good practice guides, briefings, conferences, seminars, workshops, CD-ROM for auditors, DVD of patient experiences, booklet summarising published research on a topic, self-assessment tool, articles
2006	12	Efficiency toolkit. Briefing, conferences and seminars. Manuals of good practice. Booklet on faster closing. Targeted leaflets
2007	7	Briefing for select committees, working groups, established internal practice networks
2008	5	Good practice issued
2009	8	Facilitated workshops, good practice guidance, toolkits
2010	12	Good practice guidance, guides and toolkits

Source: NAO Annual Reports.

dissemination’ – additional means of communicating findings to meet the needs of different stakeholders. The main features are:

- *Generating additional products and summaries:* briefings, guidance, informal advice and other additional products have been generated by audit teams. Some of these are specific to particular reports – such as the advice to school governors on their role sent to 22,000 schools in 2006 whilst others are more general, such as the *NAO Focus* magazine which summarised a series of reports. Others such as the ‘Short guide to structured cost reduction’, published in 2010, are designed to bring together advice on a specific issue. The rationale behind much of the outputs has been that the NAO wishes to influence those who have little time available to read detailed reports and need more accessible, summarised products, or else they are designed to provide additional, practical information for officials.

- *Using audit reports as a basis for discussion of findings:* since the 1990s the NAO has organised conferences and seminars, in recent years particularly around specific reports such as defence project management and working with the third sector, or a group of related reports such as the Learning and Innovation conference in 2009. Staff have also presented audit findings at relevant conferences. The aim of this has been to spread findings beyond the limits of those who are likely to read audit reports. Conferences have been designed to bring together front-line staff with the objective of sharing learning and stimulating practitioners to act on findings.
- *Bringing together and disseminating material on good practice:* a long running example is the electronic database of recommendations and government responses made on private finance topics (launched in 2003), to share accumulated knowledge with departments and others such as private sector contractors and advisors, linked to training workshops for public sector project teams (National Audit Office, 2004a). More recently, from 2010 the NAO website has included a separate section on 'Help for public services', which brings together all the guidance material – more than 30 products on topics such as commissioning and procurement, efficiency, financial management, project and programme management and performance measurement – in one place.
- *Developing toolkits:* in the 2000s a number of toolkits were developed and launched, aimed at public officials. These are described as drawing on NAO knowledge and being practical tools to be used by clients for self-assessment purposes, or as the basis of further NAO work. Examples include an electronic toolkit for assessing an organisation's efficiency (2007), Consultancy Assessment Tool (2007) and a decision support tool for developing financial relationships with third-sector bodies (2009), all areas examined in past VFM studies.

Our review of NAO Annual Reports leads to a number of observations. Firstly, compared with the early 1990s, there was a general increase in the following decade or so in the number of occasions on which the NAO Annual Report refers to the significance of learning or lessons from its work (although more recently the Annual Report has reduced in size and has included less material highlighting individual pieces of work, which may explain the reduction in references towards the end of the period). This can perhaps be seen as a reflection of the growing importance accorded to it by the NAO and how it would like to be seen.

Secondly, over time the NAO has become more forthright about what it sees as its wider responsibilities. In 1989, the NAO Annual Report stated

almost in passing that ‘Besides producing reports we decided in 1988 to promote improvements in value for money by identifying and publicising good management principles and practice’ (National Audit Office, 1989). By the 2000s, however, it stated clearly that ‘In addition to our financial audit and value for money work, we also have a wider role to play in advising and assisting our audit clients. Through dissemination of guidance and best practice we are able to help government departments improve their business, ultimately leading to better outcomes for citizens and taxpayers’ (National Audit Office, 2004a).

Thirdly, the range of products mentioned in NAO Annual Reports has grown in number and content. Conferences, seminars, presentations, articles and good practice guides have been seen since the 1980s, but more recently, the activities appear to involve a larger amount of effort, as well as follow-up work, requiring in some cases auditor facilitation. For example, in 2006 much of the work on efficiency related to ‘toolkits’ (National Audit Office, 2006a) which were substantial products, requiring NAO staff to lead workshops of officials. By 2010 the Private Finance recommendations database contained the accumulated knowledge from more than 80 reports, with information on over 1,000 recommendations and the associated Government’s responses. Both were clearly more substantial initiatives than simply, for example, producing journal articles. The NAO has also shown itself willing to collaborate in some cases with government bodies such as the Office of Government Commerce and the Treasury to issue guidance. Although it may be seen as a threat to its independence, it has tended to see the joint effort as strengthening the hand of each body (often at the centre of government) on areas of common concern.

The NAO’s justification for doing such work can be seen in the language employed to describe it. Producing additional outputs is clearly seen as ‘adding value’. The reasons include that it will ‘help longer-term improvements in quality of public services’, ‘prompt beneficial change’, ‘promote public service improvements’, ‘help public service managers improve performance’. And it is justified because ‘it would be a waste if we did not make the most of this knowledge, understanding and broader perspective’ (National Audit Office, 2002). For the NAO, disseminating information and insight derived from audit work has become a natural extension of its statutory role in auditing public bodies. In 2009 it commented that ‘Through our audit work, we aim to help improve the capability and performance of our clients, and help them achieve better value for money or better service provision for the citizen. We have developed a series of toolkits and good practice guides on aspects of government where we have collected a strong body of evidence which assists those delivering public services’ (National Audit Office, 2009a).

As well as secondary dissemination of lessons, there also appear to have been developments with the primary products – the published reports. A complaint in the past from government officials has been that reports were ‘too backward looking and critical’ (Lonsdale and Mayne, 2005). Over the last 10 years or so efforts appear to have been made in some reports to:

- highlight success stories and good practice examples, perhaps in boxes separate from the main text;
- produce reports specifically on success in order to highlight the factors that appear to have secured this outcome (in particular, the reports on Successful IT (National Audit Office, 2006b) and Innovation in Central Government (National Audit Office, 2009b));
- make greater use of comparisons between organisations, including drawing on examples from overseas and the private sector;
- undertake cross-government studies on aspects of policy-making such as whether joining-up is working, how successfully government is working with the third sector, the use of research in government, successful consultation with citizens (Lonsdale, 2007); and
- make increased use of the NAO website to place detailed research material, such as survey instruments and full responses.

In conclusion, the evidence presented in this section has helped to give some sense of how the NAO interprets its role in relation to learning. It appears to involve:

- disseminating information and knowledge generated from audits in a range of ways;
- creating tools designed to be utilised by front-line staff and policy makers;
- developing more constructive working relations with audited bodies and central agencies;
- adjusting the language, tone and content of some reports and incorporating examples of ‘success’;
- working with officials to assess their own performance against best practice developed from audit work and other sources; and
- promoting discussion on important issues.

What do Auditors Think Their Work Does to Help Learning in Government?

Having examined the official views, this section draws on group discussions with NAO auditors about what they considered they were trying to

achieve, what obstacles they perceived existed to learning from their work, and what makes for effective learning. Discussions were held in 2007 with 25 experienced managers and directors, who between them had over 300 years of performance audit experience.

In discussions, many emphasised that the NAO was primarily an accountability agency, a status which was considered to present constraints with regard to learning. Some argued they could operate in two different ways – doing audit or undertaking ‘improvement tasks’ – but they needed to be clear which one they were carrying out at any given time (‘You can’t be a driver, instructor and examiner at the same time’). Nevertheless, the expectation that audit went further than just helping to hold to account was widely recognised; indeed it was seen as a ‘natural development’ and ‘natural to want to help and solve problems’. Going beyond a formal role was also seen as in line with changes elsewhere in government, and ‘inevitable’ at a time when all public organisations had to justify their existence and when ‘influential outsiders’ (government officials and particularly Permanent Secretaries – the heads of government departments – were mentioned) were assumed to judge the NAO by the extent to which it contributed to helping government learn.

There was, however, some recognition of potential obstacles. Some still questioned whether auditors were well placed to offer such assistance and there was a strong belief that there was a boundary over which auditors should not step. Some suggested it would be ‘uncomfortable’, for example, to sit on a panel to implement changes in a government body and if they did they would compromise their independence and limit their ability to come back to audit the subject in the future. Reference was also made to the presence of the Committee of Public Accounts, which they felt would always influence how their work was perceived by officials. But the difficulties experienced by auditors were not considered unique to the NAO; trying to leverage change was also seen as difficult for official bodies with responsibility for driving change from the centre of government. This was because government departments were seen as highly independent and often resistant to being advised by others.

Audit was seen to be able to assist in learning in a number of ways, in particular, by producing new information and perspectives. Some considered auditors could stimulate learning simply by launching studies – an act which encouraged departments to pay considerable attention to a specific subject – or by outlining a story in a report, rather than necessarily undertaking complex analysis. In this way, audit helped to give increased attention to particular topics within a department (perhaps because the performance identified came as a surprise to senior officials), but could also raise it up the agenda of government more generally by shedding light

on a problem (the case of the NAO's hospital acquired infection report from 2000 was quoted). No reference was made to whether this might cause harm through giving greater – perhaps undue – attention to some policy fields over others. A number of those interviewed felt that government learnt 'by disasters' and that reporting on problem cases – a report on the outbreak of Foot and Mouth disease was mentioned – could help in analysing objectively what had gone wrong. At the very least, shining a spotlight on a problem was seen to generate a belief amongst senior officials that it should not be allowed to happen again. The ability of auditors to make connections between different parts of government was also seen as valuable in helping government to learn from itself because this was considered to be in contrast to the silo-based perspective of officials in individual departments.

Some auditors expressed scepticism, however, about the willingness or ability of government bodies to learn. Some doubted how much policy making was really evidence-based and whether government really had time, for example, to wait for the evidence from pilot exercises. Auditors also had misgivings about the attitude of officials to approaches from auditors to go beyond their formal accountability role. So, for example, the response to the offer of assistance with running workshops to follow up one report had been considerable suspicion about the NAO's motives.

Auditors were also asked how they developed and promulgated recommendations. Staff broadly accepted that the NAO should aim to demonstrate that recommendations would have a beneficial effect, for instance using cost-benefit analysis or a counterfactual. This was in line with the guidance issued by NAO on effective recommendations, although there were some misgivings as to how much effort could be devoted to this work. Many recognised that good client co-operation made their job of developing recommendations easier, both in diagnosing problems and identifying practical solutions, although there was also wariness at developing recommendations together where they might be easy options or favouring 'pet' approaches. However, there were different attitudes to client relations. For example, some considered it inevitable and right that the NAO sometimes had a confrontational relationship with departments. There was also some disagreement about how much detail should be given in recommendations, with many taking the view that policy makers should be left to decide how to implement fairly general assertions.

Overall, the auditors consulted recognised the ambiguous position they were in, but generally agreed that they had a role in helping organisations learn in ways beyond the formal accountability processes. Equally they were very conscious of limitations on what they could do given their status.

GOOD PRACTICE: WHAT DOES IT MEAN TO PERFORMANCE AUDITORS?

The previous section highlighted the increased production of spin-off products from audits, in particular so-called 'good practice guides' or 'toolkits'. Such products highlight either individual case studies, where in the eyes of the auditors things have been going exceptionally well, or outline generic practices of how to do things in ways which can be described as 'good' or even 'best'.

The growth in auditors promoting good practice to officials in the public sector is in line with a general surge of such advice in the UK government over the last two decades. This advice has been in many diverse forms such as official reports, brochures and leaflets, but also increasingly exclusively over the internet in online guides or good practice exchanges and databases. This surge of such forms of advice can, at least partly, be explained by the implementation of 'new public management' thinking since the mid-1990s (Overman and Boyd, 1994). This style of government is typically characterised by a move away from directive forms of control towards more non-directive forms of influencing public officials, with emphasis on benchmarking and persuasion (Pollitt, 1993; Walsh, 1995). This final section explores what role 'good practice' advice plays in particular for NAO auditors, and highlights the conditions which seem to be necessary to make such advice an effective element of evaluation and learning.

Recent research (Bechberger and Page, 2009) analysed both the prevalence and main characteristics of such advice as well as the criteria influencing its effectiveness. It showed that there has been a proliferation of advice disseminated in government and that the NAO plays a considerable role in distributing such advice in the UK public sector, with nearly 10 per cent of all advice across central government stemming from this source.¹ As already noted, the advice issued comes in various forms – separate guides or 'toolkits' – often directed at public officials (such as guidance to departments on risk management, see National Audit Office, 2004b) or sometimes even the wider public (such as guidance for school governors, see National Audit Office, 2006c). 'Good practice' has also increasingly been highlighted in audit reports themselves, often made presentationally distinct from the main text. This either takes the form of generic procedural advice (such as a list of good practice guidance for those carrying out public sector mergers, see National Audit Office, 2006d) or the brief description of specific case studies which auditors have come across during their evidence-gathering and analysis, and which they consider are worth sharing with a wider audience (such as lessons from case studies of efficiency initiatives, see National Audit Office, 2006e).

Interviews were conducted with the producers of such guidance to establish not only their motivations for producing such advice, but also to establish in more detail the evidence base underpinning it (Bechberger and Page, 2009, 7).² As with the majority of wider government good-practice advice, much of that disseminated by the NAO is not based on what would be labelled ‘rigorous scientific evaluation’ (in the sense of the authors having established a clear causal linkage between practice and purposive objective, using randomised control trials or a sampling strategy which would allow for the generalisability of results). Instead, it is often based on external expert advice or the accumulation of years of audit expertise about ‘what works’. It might also be the result of case studies which were undertaken as part of a performance audit and where auditors have come to the view that success has been sufficiently demonstrated by the auditees.

In order to assess the extent to which the kind of ‘good practice’ advice issued by the NAO fulfils the criteria for successful learning outlined earlier, it is important to stress that the production of ‘good practice’ advice by auditors fulfils quite distinct purposes. Besides the purpose of facilitating learning – which is mainly the case with advice promoting generic practices or procedures as part of audit reports – such products also appear to be a response from auditors keen not to appear too negative in their judgements of government performance and to highlight success as much as failure or shortcomings. As we saw earlier, selecting good practice case studies and highlighting them throughout a report is one way of doing this. Even if things have gone badly overall, in most instances there are at least a few examples where things went well. Finally, producing separate good practice spin-off products from their work also serves organisational objectives, as it helps those auditors keen to appear more proactive and to achieve the greatest impact from their work. Interviewees confirmed – as they did in the separate discussions referred to earlier – that they were keen to seek additional channels of influence besides the formal Committee of Public Accounts hearing route. But it could also serve a rather more self-serving purpose of appearing proactive/having an opinion in a particular area considered important.

In these circumstances, ‘methodological purity’ in establishing the causal link between recommended practice and effectiveness is not always the main concern of the producers of advice. This is particularly true when highlighting individual case studies in audit reports. A review of value-for-money reports containing such good-practice case examples shows that those are often presented without further instructions around implementation or without the auditors necessarily drawing generic, generalisable lessons from them. In most cases they leave readers to judge

whether the presented 'best' practices could be appropriate to other cases or circumstances. The methodological quality and robustness of evidence seems more important for the production of guidance material which is generic in nature and which is often published separately alongside the standard audit report, rather than the short case examples included in reports. Nevertheless, there is an important point here that auditors should be confident about the robustness of the evidence they have when they advocate particular approaches. There is also a strong argument for concentrating their 'good practice' attention on areas where they have genuine expertise.

The review by Bechberger and Page has demonstrated the potential to further improve such good practice advice in order to maximise its chances of being successful in stimulating learning. This relates in particular to a more effective dissemination of findings, as the issuing of printed and/or online guidance materials is not always accompanied by a clear strategy of how to reach and influence the target audience in the best way. If the primary purpose of issuing the guidance is to stimulate learning, improvements could be made in establishing the need for such guidance, understanding its particular timing demands, as well as the specific target audience more intensively prior to issuing the guidance.

When it comes to establishing the wider factors which make 'good practice' advice effective, the generic factors outlined in Table 13.1 are decisive. Not only the NAO-commissioned review but the wider literature around good-practice advice and policy transfer have highlighted that the adaptability of advice to specific local circumstances, its realism (in terms of the resources and time involved in its implementation), the quality of its evidence base, and the communication channels chosen to disseminate the advice (particularly via more personal channels such as professional networks) are the most important determinants of whether attention is paid to the advice and whether it is taken up by recipients (see King and Ollerearnshaw 2000; Stone 2000; Wolman and Page 2002; Mossberger and Wolman 2003; Bardach 2004; Freeman 2007).

There are indications that such lessons are having an influence on the production of good-practice advice in the NAO, where towards the end of the 2000s there was a reduction in the amount of such advice issued and efforts were made to concentrate it on particularly promising areas and audiences, particularly based around the NAO Strategy, launched in late 2009. There also seemed to be a move away from more generalist guidance to a more targeted strategy of seeking additional channels of facilitating learning, more directly aimed at Departments, which is being regarded as a more effective way of stimulating change and achieving impact from the VFM audit work.

CONCLUSIONS

This chapter set out to examine what impact working in an accountability setting had on the efforts of performance auditors to generate learning. Evidence from the literature highlighted factors which are seen to shape the ability of evaluative work to help learning. Turning to the NAO, we noted the overriding importance of its role in supporting accountability arrangements, but acknowledged that it had progressively sought to fulfil a learning role despite the limitations of its position. Reviewing Annual Reports over a 20-year period gave an indication of the initiatives that had been taken – good-practice products, hosting conferences and seminars on the back of reports and the development of toolkits and recommendations databases, and so on.

However, the ambiguity of their position came out in discussions with auditors themselves, who saw limits to how far they could assist those they audit. Auditors are often reluctant to step outside certain boundaries, and there can be resistance amongst officials to overtures from auditors, conditioned by a traditional view of audit. In looking in more detail at the preparation and dissemination of good practice material, we saw that it was often generated from examples found in the course of audit work and expert advice, rather than from scientific investigation. This raises some concerns about the credibility of the work. Against this background, we offer a number of concluding thoughts.

Firstly, it is important to emphasise that the advice and guidance that the NAO generates is a spin-off from its primary accountability role; without its statutory audit functions it would not have the knowledge to share, gained from its access and accumulated work. This is a reminder that the learning role will always be secondary. As Furubo put it in Chapter 2, if the prime purpose of the evaluative work is learning, you would not choose the format of a performance audit.

Secondly, however, it seems clear that auditors consider that both their accountability and advisory roles provide opportunities for helping with learning; they are not a case of ‘either-or’. Accountability processes have the advantage that they at least ensure lip-service is paid to implementing recommendations and considering the evidence gathered. Auditors appear to see part of their role as trying to stimulate as much learning from the accountability process as possible by doing what they can to support a constructive discussion in the formal committee hearings which consider their reports. This is alongside producing practical tools for the audited organisation and others more widely in government to use.

Thirdly, our analysis suggests that attention to learning has grown over the years and this is reinforced by the NAO Strategy, published in 2009,

which emphasises the additional value the NAO considers it can offer to those it audits. This also appears to be substantiated by the increased prominence on the NAO website of material designed to 'help public bodies'. However, setting itself up as a source of advice and guidance underlines the importance of having something worthwhile to impart. It also emphasises the need for 'good practice' and similar advice material being soundly based and credible. There is also a strong argument for focusing advice and guidance on areas of genuine expertise, rather than seeking to offer it on a too-wide range of topics.

Fourthly, it is crucial that all involved recognise that there are clear limits to how far auditors can go in assisting officials because of the importance of safeguarding their independence that Furubo highlighted as a pre-requisite for performance audit. Some of the forms of evaluation described above as conducive for learning are not appropriate for performance auditors. A blurring of the boundary between auditor and those under scrutiny cannot occur as it would undermine their credibility. They can recommend, advise, illuminate and facilitate discussions, but they cannot implement.

NOTES

1. A total number of 951 pieces of 'good practice' advice were identified and included in the analysis; this was based on a systematic web-based search of all published documents and materials by UK government departments as well as the NAO, Audit Commission and Office of Government Commerce between January 2004 and March 2007. 'Good practice' advice was defined as documents which (a) made a recommendation, (b) referred to an identifiable set of practices, (c) adhered to the principle that advice is voluntary, and (d) were generalised, i.e. given to people within different organisations or locations.
2. Interviews were conducted with seven NAO teams who produced 'good practice' advice as part of their Value for Money reports.

REFERENCES

- Bardach, Eugene (2004), 'Presidential address - The extrapolation problem: How can we learn from the experience of others?', *Journal of Policy Analysis and Management* **23** (2): 205–20.
- Bechberger, E. and E. Page (2009), 'Simply the best? Understanding the market for "good practice" advice from government research and evaluation', Report to the National Audit Office, London School of Economics, London.
- Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale and Burt Perrin (2007), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Jersey: Transaction.

- Black, Robert (2000), *Supporting Democratic Scrutiny by Public Audit*, London: Public Management and Policy Association.
- Bourn, John (2007), *Public Sector Auditing: Is it Value for Money?*, Chichester: Wiley and Sons.
- Bovens (2004), 'Public accountability', in Ferlie, E., L. Lynne and C. Pollitt (eds), *The Oxford Handbook of Public Management*, Oxford: Oxford University Press.
- Bowerman, M. (1996), 'The rise and fall of VFM auditing', in Lapsely, I. and F. Mitchell (eds), *Accounting and Performance Measurement: Issues in the Private and Public Sectors*, London: Paul Chapman Publishing, pp. 193–212.
- Bowerman, M., C. Humphrey and D. Owen (2003), 'Struggling for supremacy: The case of UK public institutions', *Critical Perspectives on Accounting*, **14**, 1–22.
- Brazier, A. and V. Ram (2006), *The Fiscal Maze: Parliament, Government and Public Money*, London: Hansard Society.
- Freeman, Richard (2007), 'Epistemological bricolage: How practitioners make sense of learning', *Administration and Society*, **39** (4), 476–96.
- Gendron, Y., D. Cooper and B. Townley (2001), 'In the name of accountability: State auditing, independence and New Public Management', *Accounting, Auditing and Accountability Journal*, **14** (3), 278–310.
- Gendron, Y., D. Cooper, and B. Townley (2007), 'The construction of auditing expertise in measuring government performance', *Accounting, Organizations and Society*, **32** (2007), 101–29.
- Howlett, M. and M. Ramesh (1995), *Studying Public Policy*, Oxford University Press.
- King, E. and S. Ollerearnshaw (2000), *The Effectiveness of Different Mechanisms for Spreading Best Practice*, London: Cabinet Office.
- Landers, B. (1999), 'Encounters with the public accounts committee: A personal memoir', *Public Administration*, **77** (1), 195–213.
- Lehtonen, M. (2005), 'Accounting (f)or learning?', *Evaluation*, **11** (2), 169–88.
- Lipsey, D. (2000), *The Secret Treasury*, London, Viking.
- Lonsdale, J. and J. Mayne (2005), 'Neat and tidy . . . and 100% correct': Assuring the quality of Supreme Audit Institution performance audit work', in J. Mayne and Schwartz, R. (eds), *Quality Matters: Seeking Confidence in Evaluating, Auditing and Performance Reporting*, Transaction, New Jersey.
- Lonsdale, J. (2007), 'Helping government to be responsive, responsible and respected: A reaction to Jocelyne Bourgon', *International Review of Administrative Sciences*, **73** (1), 31–6.
- Martin, S. (2005), 'Evaluation, inspection and the improvement agenda', *Evaluation*, **11** (4), 496–504.
- Massey, A. and R. Pyper (2005), *Public Management and Modernisation in Britain*, Basingstoke: Palgrave Macmillan.
- Mossberger, K. and H. Wolman (2003), 'Policy transfer as a form of prospective policy evaluation: Challenges and recommendations', *Public Administration Review*, **63** (4), 428–40.
- National Audit Office (1989), *Annual Report*, London: National Audit Office.
- National Audit Office (2002), *Annual Report*, London: National Audit Office.
- National Audit Office (2004a), *Annual Report*, London: National Audit Office.
- National Audit Office (2004b), *Good Practice in the Application of Risk Management – Self Assessment Questions for Departments*, London: National Audit Office.
- National Audit Office (2006a), *Annual Report*, London: National Audit Office.

- National Audit Office (2006b), *Delivering Successful IT-Enabled Business Change*, London: National Audit Office.
- National Audit Office (2006c), *Improving School Performance: A Guide for School Governors*, London: National Audit Office.
- National Audit Office (2006d), *The Creation of Ofcom: Wider Lessons for Public Sector Mergers of Regulatory Agencies*, London: National Audit Office.
- National Audit Office (2006e), *Progress in Improving Government Efficiency: Lessons from Case Studies of Efficiency Initiatives*, London: National Audit Office.
- National Audit Office (2009a), *Annual Report*, London: National Audit Office.
- National Audit Office (2009b), *Innovation across Central Government*, London: National Audit Office.
- National Audit Office (2010), *Annual Report*, London: National Audit Office.
- OECD (2001), *Background Synthesis Report*, Paris: OECD.
- Overman, S. and K. Boyd (1994), 'Best practice research and postbureaucratic reform', *Journal of Public Administration Research and Theory*, **4** (1), 67–83.
- Pollitt, Christopher (1993), *Managerialism and the Public Service: The Anglo-American Experience*, 2nd edition, Oxford: Blackwell.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilka Summa and Marit Waerness (1999), *Performance or Compliance? Performance Audit and Public Management in Five Countries*, Oxford: Oxford University Press.
- Preskill, H. and R. Torres (2000), 'The learning dimension of evaluation use', in *The Expanding Scope of Evaluation Use*, (eds) Caracelli, V. and H. Preskill, *New Directions for Evaluation*, Number 88, Winter, San Francisco: Jossey-Bass Publishers.
- Rossi, P., H. Freeman and M. Lipsey (1999), *Evaluation: A Systematic Approach*, Thousand Oaks, California: Sage Publications.
- Sharma, N. (2007), 'Interaction and interrogations: Negotiating and performing value for money reports', *Financial Accountability and Management*, **23** (3), August.
- Stone, Diane (2000), 'Non-governmental policy transfer: The strategies of independent policy institutes', *Governance*, **13** (1), 45–70.
- Taut, S. and D. Brauns (2003), 'Resistance to evaluation: A psychological perspective', *Evaluation*, **9** (3), 247–64.
- Van der Knaap, P. (2006), 'Responsive evaluation and performance management: Overcoming the downsides of policy objectives and performance indicators', *Evaluation*, **12** (3), 278–93.
- Van der Meer, F.-B. and J. Edelenbos (2006), 'Evaluation in multi-actor policy processes', *Evaluation*, **12** (2), 201–18.
- Walsh, K. (1995), *Public Services and Market Mechanisms: Competition, Contracting and the New Public Management*, London: Macmillan Press Ltd.
- Weiss, Carol (1998), *Evaluation*, New Jersey, Prentice-Hall, 2nd edition.
- Wilkins, P. and J. Lonsdale (2007), 'Public sector auditing for accountability: New directions, new tricks?', in Bemelmans-Videc, M.-L., J. Lonsdale, and B. Perrin, (eds) *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Jersey: Transaction.
- Wolman, H. and E. Page (2002), 'Policy diffusion among local governments: An information-theory approach', *Governance*, **15** (4), 477–501.

14. Responsiveness in performance auditing: towards the best of both worlds

Peter van der Knaap

Progress and practices must be built on learning from experience.
(INTOSAI 2010, 29)

INTRODUCTION

There is a natural tension between performance auditing on the one hand, and the complex and dynamic nature of public policy on the other. As Vital Put discussed in Chapter 4, performance auditors tend to work on the basis of professional audit norms and criteria that form the basis for performance judgements. Increasingly, they make use of policy objectives, targets, and indicators. Combined, these norms, criteria, policy objectives and indicators tend to become the stable, rather static, frame of reference for the performance auditor. Yet in the complex and ever-changing world in which public policy makers operate, circumstances, knowledge and preferences continuously change.

Responsive and responsible policy making is about keeping track of these dynamics, and what we might call ‘responsive performance auditing’ should contribute to this. Such work entails paying close attention to changes in context, knowledge, and stakeholders’ beliefs and preferences, and deciding how to incorporate these changes into performance judgements, conclusions and recommendations for improvement. Responsive performance auditing does not mean that the auditor should abandon traditional criteria for judgement and become merely a passive recipient of other people’s interpretations and criteria. Instead, it should enhance the auditor’s repertoire by adding deliberate consideration of dynamic complexity and change to more static audit procedures, methods, and standards.

How can we get the best of both worlds? How can we conduct performance audits so that, on the one hand, we maintain the benefits that

clear audit standards, norms and criteria, together with the focus that policy objectives and performance criteria provide, while at the same time keeping an eye open for its negative consequences and the desirability of adapting to changing contexts, evolving knowledge, and developments in preferences?

In this chapter, developments towards a more responsive performance audit approach are explored. It describes the debate on the complex and dynamic nature of public policy, and some of the risks of 'static' public policy management and performance auditing are described. As a pointer towards greater responsiveness, recent developments in the audit practice of the *Algemene Rekenkamer* (Netherlands Court of Audit or NCA) are described. Finally, some conclusions are presented, highlighting how, in changing environments, performance auditing can contribute to policy-oriented learning by taking a dynamic, rather than a static approach, and focussing primarily on results, rather than on process.

PERFORMANCE AUDITING: DEVELOPING TOWARDS GREATER RESPONSIVENESS

There is a classic episode from the 1980s British television comedy 'Yes Minister' in which Sir Humphrey – a senior civil servant – meets for a drink with another senior official – Sir Ian Whitchurch, head of the Department of Health. The latter expresses surprise at Sir Humphrey's minister's concern over the situation at a particular hospital. As he explains, it has no patients because there are no nurses, and he argues, one cannot set up a hospital and start it running smoothly when there are patients around. He tells Sir Humphrey to inform his minister that this is the 'run-in period'. The case for a patient-free hospital is strengthened, however, by its superb performance on staff satisfaction and overall quality, measured by the absence of accidents and of complaints.

The episode perfectly illustrates one of the main pitfalls of what can be called traditional, static performance auditing. As critics point out, the default mode of both performance auditors and those auditing financial statements may be to fervently stress the importance of objectives, targets and criteria, and criticise shortcomings in this domain, without bothering too much about whether those performance indicators still relate to society's real problems, or are actually in need of revision (WRR, 2007). Yet it is precisely here that responsive performance auditing can contribute.

Responsive performance auditing can be defined as a type of performance auditing in which the auditor still works from a basis of well-established standards, norms, and criteria, but at the same time:

- (1) deliberately investigates whether changes in contexts, knowledge, and preferences have occurred; and
- (2) weighs up whether the auditee responded well to these changes before final judgements are made.

Part of performance auditing, therefore, is the employment of ‘dynamic norms’. Learning, and the way it is organised, is crucial in this respect. How do governments and agencies know what works best in which conditions? How do they organise their own feedback through monitoring and evaluation? How do they consult with stakeholders? And, crucially, how do they translate the resulting lessons into better practice?

As we saw in Furubo’s chapter, audit is generally defined as an objective assessment of either compliance with applicable statutes and regulations (regularity audit) or of economy, efficiency and effectiveness with which policy programmes have been implemented. Traditionally, performance audit is defined as an independent examination of the efficiency and effectiveness of government undertakings, programmes or organisations, with due regard to economy, and the aim of leading to improvements (INTOSAI, 2010, 11). Often, questions related to economy, efficiency, and effectiveness pertain to rather formal or indeed static elements of policy programmes. The INTOSAI guidelines, for instance, stress that performance auditing is based on decisions made or goals established by the legislature. Typical questions of interest to a performance auditor would be: is there a clear structure of performance goals and have the appropriate priorities and instruments been chosen for the use of public funds? Is there a clear distribution of responsibility between the different levels of authority? Is there general cost awareness? Is there an adequate emphasis on management controls and reporting requirements? (INTOSAI, 2010, 12).

However, the need for a more responsive way of auditing is also increasingly acknowledged. It should be noted that the INTOSAI guidelines include auditing the ‘actual impact’ of activities compared with the intended impact as part of performance auditing (INTOSAI, 2010), and an orientation towards citizens’ needs is advocated. In addition, the guidelines specifically state that:

While financial auditing tends to apply relatively fixed standards, performance auditing is more flexible in its choice of subjects, audit objects, methods, and opinions. Performance auditing is not a regular audit with formalised opinions, and it does not have its roots in private auditing. It is an independent examination made on a non-recurring basis. It is by nature wide-ranging and open to judgments and interpretations. (. . .) It is not a checklist-based form of auditing. The special feature of performance auditing is due to the variety and complexity of questions relating to its work. (INTOSAI, 2010,12)

Unlike compliance auditing, performance auditors should not just look at whether policy programmes are being carried out according to plan. Neither should they focus too narrowly on whether objectives are reached, criteria met, and indicators attained. Two simple questions help the performance auditor in selecting the audit angle: (1) 'are things being done right?' and (2) 'are the right things being done?' (compare Booth and Evans, 2006). The first question is primarily aimed at what may be called the 'producer', and is concerned with whether policy decisions are being carried out properly. Here, the regularity auditor will want to know whether the executive has observed the rules or the requirements, whereas performance auditing has been mainly concerned with different aspects of the economy and efficiency of the operations (INTOSAI, 2010).

The scope for analysis becomes considerably wider, however, when the second question – whether the right things are being done – is asked. First, this question refers to effectiveness: the impact of policy measures on society. Here, the degree to which measures lead to the desired outcome (that is, the realisation of the objective) is to the fore. The conclusion of a performance audit or evaluation may be that a certain measure proved to be ineffective or to be inconsistent with other, more important, objectives. Effectiveness is of course still a very important criterion: the question whether or not government or public agencies 'perform' or 'deliver' must be related to the results it achieves or helps to achieve. Earlier formulated objectives are key, both from a democratic government perspective and a more managerial accountability perspective at the level of agencies. More fundamentally, however, the question, 'are the right things being done?' implies that any government undertaking, any programme or measure to achieve a certain objective, is open to debate. Despite the effectiveness of programmes, the question can be: do we really want this outcome? Are we still considering it a priority on which we want to spend public money?

PUBLIC POLICY AS AN AUDIT TOPIC: COMPLEXITY AND DYNAMICS

In 2002, the Dutch sociologist-turned-politician Pim Fortuyn accused a campaigning fellow politician of lack of professionalism when she exclaimed that some problems could not be solved easily by policy measures due to the 'complexity of the problem and the dynamic nature of its context.' Despite this – and Fortuyn knew this very well – many of modern society's problems are indeed complex, and change is the only constant factor.

This perspective is echoed in a report by the Scientific Council for Government Policy in the Netherlands on 'the learning government.' The

Council argues that many complex and dynamic problems facing our society need to be at least ‘tamed’ before we can understand and define them in clear terms (WRR, 2007). This ‘defining’ is also necessary to make possible solutions part of formal structures of democratic decision-making and, thereafter, implementation. The Scientific Council calls these complex and dynamic problems ‘untamed’: characterised by being ‘constituted by a high degree of cognitive and normative uncertainty.’ According to the council, taming these problems entails more than democratic deliberation and negotiation between parties; it takes a much wider approach including consultation, local experiments, comparative assessments and research.

Audits are not mentioned by the Council, but if performance audits are to contribute to real policy results, they must also pay close attention to the complex and dynamic nature of public policy. There are several factors that determine this characteristic. Some of these factors may be labelled as autonomous. Modern western societies have substantial normative differences between groups, combined with different centres of power and influence that overlap and conflict from domain to domain. Diversity and complexity are added to by the expansion of information, new technologies and multiculturalism. Governments can use different instruments to increase openness to these dynamics such as policy evaluation, consultations, city hall meetings, and experts and citizens panels (potentially using web 2.0 applications).

Other sources of complexity are not autonomous in themselves, but rather, they are the consequences of choices about how to deal with complexity (Frissen, 1996). These choices – some made more consciously than others, with many actually a part of our cultural heritage – pertain to the way political power is organised, public policy is implemented and markets are regulated (Frissen, 2007). Furthermore, many of the complex relations between central government and local administrations, or the implementation of measures by quangos, agencies and other administrative bodies are man-made.

As the latter complexity can be remedied more easily (but is – at the same time – often harder to detect), it is important to recognise the difference between on the one hand, autonomous and, on the other, man-made complexity and dynamics. Performance audit may help in this respect. One way for auditors to contribute to a more knowledge-oriented public policy domain that has better judgement capacity is to actively promote policy experiments (pilots, test programmes) and research initiatives that are specifically designed to policy-oriented learning. In addition, it is helpful to be aware of the ‘taming trap’ that well-articulated objectives, indicators, and criteria bring with them. This is that problems appear ‘under control’, while important negative consequences may be overlooked.

PERFORMANCE MANAGEMENT: ADVANTAGES AND RISKS OF 'STATIC AMBITIONS' IN OBJECTIVES, INDICATORS, AND CRITERIA

It is naturally tempting to think of society's problems as 'controllable' or 'tameable' and of policy as developing in sequence: it makes things simple and allows for clear statements, enticing political campaigns, and a clear division between policy and implementation. In the light of this taming and sequencing, complex and ever-changing problems can be portrayed as having fixed ambitions. Hereafter, politicians and public sector managers – and performance auditors – can focus comfortably on objectives, targets, and criteria.

Such an 'objective-and-indicators-driven' approach may bring advantages, but it also has risks. Objectives, targets, and criteria may lead to simplicity, resistance to change, and unwanted strategic behaviour (gaming) of the kind that Leeuw discussed earlier. Objectives, targets, criteria, and performance indicators are – by nature – logical abstractions of reality. They represent a simplified version of reality, and reflect the normative ideal 'that monitoring systems ought to replace the complex social-political processes entailed in the design and delivery of social and educational services' (Schwandt, 2002: 9). Schwandt's argument is clear: performance indicators can never do that and will even distort the understanding that we develop. There is a risk that the use of objectives, targets, and indicators may result in not seeing unexpected effects, changing circumstances or public preferences.

In addition, policy implementation is often a long process. In many cases, government is not the dominant actor (and even if it is dominant, it is not all-powerful). Especially when applied in 'steering relationships' and accountability systems, strategic behaviour will inevitably lead to learning processes and strategic behavioural responses (In't Veld, 1989: 28). The use of objectives, criteria, and indicators may lead to perverse effects.

As an example, by presenting a large number of policy targets and performance indicators, agents may effectively obscure the real issues at stake. In some policy domains the phenomenon of 'data dumping' leads to bulky and often incomprehensible documents. As one review put it, 'True insight into the key questions is still too often frustrated by the natural tendency of officials and administrators to hedge their bets and give veiled answers. After all, it saves them problems. But it is definitely an obstacle for the Lower House to do its work properly' (Ministry of Finance, 2004:17).

Bovens depicts this phenomenon as the 'blame game': presentational, policy, or agency strategies to minimise or avoid blame in case of failures, and to maximise credits for successes (Bovens, 2005, 2009). Because of

this gaming, authors like Behn and Power have pointed to the tension between accountability and efficient government. Too much emphasis on accountability and, with it, the importance of policy targets and performance indicators, can lead to sub-optimal and inefficient decisions instead of improved performance. In the Netherlands, this risk was stressed by an influential report by the Government Scientific Council on the 'Proofs of good service provisions', which argued 'focusing exclusively on measurable performance will lead to a destruction of knowledge, a "dictatorship of mediocrity" and a "stacking-up of supervision and responsibility"' (WRR, 2004: 4).

It is important to point out that this pitfall is not remedied by most evaluations. Moreover, a similar effect may even occur if evaluations are tailored too strictly to existing policy targets and/or performance indicators. As a result, the evaluation research may focus exclusively upon the fixed policy and underlying policy theory, its goals and performance indicators (Van der Knaap, 2006). In this respect, the 2004 evaluation of the budget reform in the Netherlands found that, in general, evaluation reports made no attempt to compare recent developments in policy priorities with long-term objectives and trends as formulated in the policy agenda. It noted 'Questions such as "Are we on the right course?" are rarely asked' (Ministry of Finance, 2004:19). In addition, the basic question of why government is at all involved in certain issues is not addressed by evaluations. According to the report, this 'bias towards conservatism' may be explained by the fact that there is lack of independence in the way evaluations are commissioned and/or carried out.

Despite this, using objectives and indicators can provide focus. They help to reduce complexity, which enables policy makers to devote their attention to the 'most important issues' and to develop a clear, communicable vision on how to achieve what they are aiming for. Objectives and indicators can be used to create a frame of reference in a policy-oriented debate. They follow the rationality of the influential results-oriented management paradigm, in which expectations on the policy measures and their effects set the stage for managing, measuring, and judging efficiency and effectiveness. When agreement is reached on objectives and indicators, they form the starting point of collective action on the basis of which interim adjustments may be made and learning can take place. The advantages and risks of objectives and indicators are set out in Table 14.1 (overleaf).

Performance audit has an important role in optimising the benefits of objectives and indicators while, at the same time, contributing to dealing with the challenges and traps they pose in an intelligent, transparent way.

Table 14.1 Advantages and risks of objectives and indicators

	Benefits of policy objectives and performance indicators:	Challenges:
Policy development	Focus	Simplification, tunnel vision and rigidity
Policy-oriented debates	Framework of reference	Simplification, tunnel vision and rigidity plus closure
Implementation	Deliberate choice between improvement or innovation, judging success or contents	Wrong level learning: conservatism or 'innovation'
Evaluation and accountability	Relevance, connection	Framing, defensive routines

RESPONSIVENESS IN POLICY EVALUATION AND PERFORMANCE AUDIT LITERATURE

Although as we have seen performance audit and policy evaluation have their differences, important lessons can be drawn from policy evaluation literature (see Bemelmans-Videc, 2007). Within this literature, there is a pendulum movement in which the traditional, rational-objectivist model of policy evaluation is repeatedly challenged and defended (see Majone, 1989; Dryzek, 1993; Fisher and Forester, 1993; Van der Knaap, 1995).

In the last decades, a more responsive, argumentative approach has been advocated in the evaluation literature, in which evaluation would be geared less to formal objectives and indicators (policy intentions), and more to actual activities, attitudes, and 'real results'. In this way, evaluators would be less determined by objectives and indicators in their assessment of outcomes as 'successful' or 'disappointing'. Instead, the researcher could be more responsive to actual changes in perceptions, attitude, and/or behaviour. When reporting on the results of policy programmes, the evaluator would explicitly refer to the stakeholders' opinions, values and standards, especially when they differ from official policy documents (Abma, 1996: 63; Van der Knaap, 2006).

Authors like Guba and Lincoln (1989), Abma (1996), and Schwandt (2001, 2002) stress the pluralist, interdependent nature of both government and society. In line with Frissen and the Scientific Council WRR, they argue that the successful development and implementation of public policies requires the support, participation, and often cooperation of many actors. In addition, they witness an ever-increasing complexity of modern society's problems. Central to the argumentative-responsive

approach is the belief that, through constructive argumentation, policy actors, networks or advocacy coalitions may arrive at better judgements on policy issues and, hopefully, at 'better' policies and ways of delivering those policies.

I would argue that this kind of responsiveness to social change and changes in stakeholders' perspectives is crucial for any political system that wants to intervene in society's problems in an intelligent, informed, and well-balanced way (Van der Knaap, 2006). This is especially so when applied in a system of results-oriented budgeting, where policy objectives and performance indicators play important roles, and the added value of policy evaluation should principally be the systematisation of the possibility that dialogue will reveal differences (Majone, 1989).

In evaluation literature, various approaches that can bring more responsiveness into evaluations are put forward. Perhaps the most famous one is the goal-free evaluation model. Herein, 'evaluation criteria to assess the program's effectiveness are not only derived from the goals and intentions of policy makers, but include a wide range of issues of as many stakeholders as possible' (Abma, 2005:279; compare Stake, 2004). In short: the evaluation tries to establish the actual effects of a policy programme, not just those adopted by a central actor (Scriven, 1991). Observing without preconceived ideas about the outcome of a programme, let alone checklists, is the key to this approach. Description and an independent assessment by the evaluator of the importance of what she or he comes across are central to this model. Responsive evaluation actively seeks to enrich a policy debate by bringing in alternative perspectives, new explanations for success or failure, and 'emancipating minority viewpoints'. The main challenge lies in demonstrating the validity of responsive methods and developing the skills that this kind of research requires (interpersonal relationships, communication).

Despite continuous debate on the merits and methodological validity of responsive evaluation, it is broadly accepted that it has added value to results-oriented ways of developing and delivering public-policy programmes. However, there are differences between evaluation and audit. As Ling puts it, evaluators may be able to live with 'multiple truths', and accept the notion that different stakeholders have different views on the performance of a programme (Ling, 2007, 133). For auditors, this is often different but in the audit literature, too, due attention is given to responsiveness. Along with INTOSAI, there are others who point at the need to recognise the complexity of public-sector interventions. As Bemelmans-Videc et al. note: 'rarely does a program or intervention work completely independently of the interventions of other actors or the influence of other factors' (Bemelmans-Videc, Lonsdale and Perrin, 2007, 248). The authors

argue that in changing environments, ‘responsive programs should be changing their objectives to take advantage of opportunities or to respond to feedback [including feedback from audit and evaluation]’ (ibid., 249). They envisage a new type of accountability in which public sector auditing should play its role by (1) taking a dynamic rather than a static approach, (2) focussing primarily on results rather than on process, (3) stressing the need for continuous and responsive learning (ibid., p.253; compare also Van der Knaap, 2007). How this perspective, which this author strongly supports, has already been applied in performance auditing, is the topic of the next section.

TOWARDS MORE RESPONSIVENESS IN PERFORMANCE AUDITS: NETHERLANDS COURT OF AUDIT

As noted above, the need for a more responsive approach to auditing is increasingly acknowledged (for example, INTOSAI, 2010). Audit institutions like the Netherlands Court of Audit (NCA) and the National Audit Office in the United Kingdom have already tried to take a more responsive approach and stress the need for learning. To the NCA, there are three dimensions to this:

- the selection of audit themes and topics: putting social relevance and risks at the centre of choices (see the chapter by Put and Turksema);
- the audit process: giving auditees and stakeholders a say in the process, using more participatory methods; and
- follow up reports: systematically keeping track of the use and usefulness of performance audits.

Responsiveness in the NCA Audit Process

In its audits, the Netherlands Court of Audit tries to be open and responsive, both internally and towards the individuals and organisations it deals with. The first aim is to be transparent about the background and nature of performance audits, the methodological approach, and the audit standards and criteria. By doing so, the NCA intends to prevent unwelcome surprises for auditees and other stakeholders (Court of Audit, 2005).

Being open about audit intentions and approach and being open to criticism is, however, only the first step. For effective and relevant audit results, auditors must be responsive to what auditees and other stakeholders have to say about the relevance of audit questions, the meaning of its

intermediate findings, and the usefulness of potential recommendations.¹ In 2005, under the heading of 'the effective NCA', the NCA has actively opted for a more open and responsive approach, including more participatory methods. The NCA has developed a 'menu for effective performance audits' (Court of Audit, 2005). This menu presents different audit methods and activities that audit teams can use during the different stages of the audit process. It is based upon an inventory of such techniques which have been applied effectively by NCA audit teams in the past.

The responsive audit approach consists of two phases: (1) 'actor' analysis and (2) selection of possible methods and actions. In the actor analysis phase, the NCA auditors set out to analyse the main actors and stakeholders, and their respective interests. The auditors look into the main parties involved, their involvement in a policy programme, recent communications and possible perceptions of the upcoming audit. In addition, the sensitivity of the topic in question, the complexity of the policy domain or network, and the level of agreement on problem definitions, relevant values and norms, and the necessity for improvement are considered. The result of this first phase is a picture of the auditee and, if relevant, other stakeholders, along two dimensions: (a) the level of trust, ranging from high to low, and (b) the reciprocity of interests, ranging from mutual to opposite. If there is a high level of trust and interests are shared, a more cooperative approach may be fruitful: lessons and insights may lead to early adaption by a ministry or even stakeholders. If, on the other hand, interests differ and there seems to be little trust between the auditor and auditee, a more formal approach may be better, while the use of external experts may help to obtain specific insights or specialist feedback.

In the second phase, the audit team makes a choice of possible actions. Along the timeline of the audit process, actions such as meetings at the start of the audit and brainstorming sessions to discuss intermediate outcomes and potential remedies are considered and plotted. An important example of these responsive methods is the organisation of expert-panels in which auditees and other stakeholders are invited to discuss intermediate audit findings and possible directions for policy programme or implementation improvement.

For example, for the NCA's audit of the energy savings policy programme, a panel session was organised with auditees from the Ministry of Economic Affairs and the Ministry for the Environment. In addition, representatives from the paper and plastics industry, research and university institutions, and an environmentalist awareness organisation were invited. The NCA team presented their intermediate audit findings, shared their dilemmas on what to conclude and possibly recommend, and tested some

of their assumptions on what would work better in which circumstances. In another audit of the Ecological Structure in the Netherlands, the NCA auditors did the same with public officials from the ministries involved, and stakeholders from the forestry, nature, and farming sectors.

‘Reality Checks’ by the NCA

Shortly after ‘Accountability Day’ 2009,² the House of Representatives passed the Pechtold motion to improve ‘auditability’ and accountability for the achievement of the government’s goals. This prompted the NCA to explore a new approach. In 2009 and 2010, the Netherlands Court of Audit carried out 20 ‘reality checks’. The NCA deliberately used this – rather provocative – term in order to stress the ‘checking’ dimension of policy objectives and performance indicators. As governments and agencies tend to develop what we described earlier as ‘frozen ambitions’ in isolation from others, the Court wanted to stress the importance of the perspective of citizens and businesses (small and medium-sized enterprises) as a starting point and frame of reference for its audits. In the words of the NCA President when presenting the results to parliament: ‘With our reality checks the NCA wants to bring the perspective of citizens and businesses to the “Binnenhof”’ (the seat of central government in The Hague).

The pivotal question was: what is the contribution made by 20 public policy programmes to resolving social problems from the position of the ‘policy addressee’ – the person, business or institution that is experiencing the particular problem? What concrete impact does the policy have for those directly concerned? The NCA made an assessment across four dimensions, each revealing how government policy contributes to resolving social problems. These were:

1. the perceived relevance of the policy programme;
2. the appraisal by target groups of its implementation;
3. the degree to which spending actually ‘landed on target’ (the money was being spent on achieving the government’s goals); and
4. the information on the effectiveness of the policy measures taken.

For each of the 20 audits the NCA started out by looking at the societal problems that affect citizens and businesses, and which government aims to solve or contain. The auditors investigated the ways in which the state had translated these problems, and how they were perceived, into policy objectives and measures. They ‘checked’ with the recipients of policy measures how successful this was and how the implementation was judged. In

BOX 14.1 STIMULATING RAIL TRANSPORT

The policy to increase the number of miles travelled by passengers on the public rail system in the Netherlands was established after a broad social consultation. Elements mentioned by travellers were included in the programme (for example, better signposting).

In addition, the audit team looked at the degree to which funds were actually spent on the designated objectives and what was known about the effects of the measures.

Most of the 20 policy programs the NCA investigated were considered to be relevant and focussed on real problems by those on the ‘receiving end’ of measures. Moreover, the auditors found that Dutch central government often explicitly invested in responsive action by measuring and monitoring preferences amongst target groups before embarking on new policy programmes (see Box 14.1). In addition, reports by central planning agencies and intermediate evaluations were used to make sure the policy measures were, and remained, ‘in touch’ with the issues at stake and the concerns of those affected.

The audit also found, however, that there often exist multiple layers between central government and the end recipients of public policy measures. For example, for both gathering information on what is needed and the actual delivery of policy measures, central government depends on other organisations, notably local government and quasi-autonomous governmental bodies (‘quangos’). The risk of such interdependencies can be an increasing distance between government and policy recipients. This may be added to by governance demands for steering and accountability purposes: system and institutional rationality and relevance may come to crowd out societal relevance and rationality. On the other hand, it often makes sense to employ these intermediate organisations, as they may be better positioned to recognise specific, context-bound circumstances that surround societal problems.

In a few cases, the NCA came across a ‘shower-of-shot-approach’. Instead of evidence-based interventions, with a clear, well-established relationship between intervention and effect, there were exceptions where a rather rash choice of subsidies was made: funds that are put at the disposal of citizens, businesses or intermediate organisations. ‘Money will work wonders’ seemed to be the dominant perspective, even when other factors (cultural, technical, judicial) contributed to the problems.

BOX 14.2 CHILD PARTICIPATION

The policy to increase the participation of children in sporting and other organised activities is carried out by the municipalities. Central government, which transfers the budgets to local government, has no say in the actual projects, but it can and does make agreements to stimulate municipalities to give families 'tangible help' (for example: paying back sporting club's membership fees).

BOX 14.3 INCREASING THE NUMBER OF FEMALE PROFESSORS AT UNIVERSITIES

The aim of the measures to increase the number of female professors at universities is to, ultimately, contribute to the quality of university research and academic teaching. Although it is possible to measure the percentage of female professorships, this is not the only factor in stimulating academic excellence.

In the majority of the 20 reality checks there was a clear relationship between the means deployed and the performance rendered in terms of services and products delivered. In some cases, however, this relationship was not clear. The interdependency of central government and local agencies also had implications for the way money was spent in relation to policy objectives. The NCA found that, in dealing with local government and other decentralised or specialised agencies, there are often 'open target spending' arrangements (see Box 14.2). This involves the implementing agency being free – to some extent – to choose where to spend the money.

In most public policy domains, the link between cause and real societal effects is difficult to establish (see Box 14.3). This is reflected in the way central government accounts for its effectiveness; rarely does the annual account of a ministry provide a clear and useful account of the relationship between policy measures and effects, let alone the way in which it reacted to disappointing results. Instead, intermediate effects are chosen.

The NCA launched the reality checks as an innovative project, in which it deliberately put responsiveness as the guiding principle of its audit approach. In response, the Ministry of Finance (which has introduced responsiveness into its own work as Box 14.4 illustrates) recognised the

BOX 14.4 MINISTRY OF FINANCE'S POLICY AUDITS

The Netherlands Ministry of Finance also actively seeks to build-in responsiveness in its performance audits. A new instrument – the so-called ‘policy audits’, which have been carried out since 2006 by the Ministry across the whole of the Netherlands central government – is a good example of this ambition. The policy audits focus on the objectives of budget-line items. They are carried out periodically, and pay specific attention to the underlying problem analysis, the actual relevance of the policy objectives, and the reason for government intervention. Questions include:

- What is the problem that gives (has given) cause to the policy? Is this problem still a topical subject?
- What is the cause of the problem?
- Why does the government consider that it is its responsibility to solve the problem?
- Why is the responsibility at national government level (and not at local or EU level)? How is the responsibility organised and why?
- Which policy instruments are used? What is the connection between the instruments? Is there any overlap? What is known about the implementation of the policy and the efficiency of operational management?
- What is the effect of the instruments on the formulated objectives (problem solution)? Do the instruments have important effects on the formulated objectives of other policy areas? What are important positive and negative side effects?

importance of insight into the impact that policy has on resolving social problems.

FOLLOWING-UP ON NCA AUDITS THROUGH IMPACT ASSESSMENTS

What happens with audit recommendations and the ministerial undertakings after audit reports have been published? Were recommendations

followed and did that really help? For any performance auditor, these are of course important questions, as the earlier chapters by Funkhouser and Weets have shown. To find the answers, the NCA actively monitors the ministries' follow-up to its audit recommendations, including, if necessary, for five years or longer.

For example, in 2006, the NCA published an audit report on 'Jobless people without a basic qualification'. The audit investigated central government policy to give jobless people aged 23 and over without basic qualifications better prospects of finding a job. Three years later, in 2009, an impact assessment was carried out. The NCA focussed on just three recommendations that went to the heart of the problem, which was the poor organisation of the long-term reintegration of jobless people with no qualifications by means of training and education. The NCA found that a lot had been achieved to promote regional cooperation in the fields of education, employment and long-term reintegration. The cooperation was directed chiefly at courses that combined work and learning, with less attention paid to the sustainability of the cooperation. Courses that combined work and learning are designed for both those in work and the jobless, and are not reserved exclusively for jobless people without basic qualifications. Reaching a specific group calls for a targeted approach.

The two questions considered by the auditors were, firstly, 'Were recommendations followed?', and secondly, 'Did they really help?'. These indicate that follow-up audits have two aspects. The first is rather straightforward: did the ministries do what they promised to do? From the example above, the NCA concluded that regional cooperation to help jobless people without basic qualifications had been strengthened. In the second aspect, the NCA investigated which changes – positive and negative – actually occurred in relation to the issue at stake. Here, much harder questions are asked: how did the follow-up of audit recommendations contribute to the quality of implementation processes? Were jobless people better helped? To what degree was the social problem solved? Did they succeed in finding new employment?

The impact assessment forms an important instrument for the NCA's performance auditing to remain responsive. In the example, the NCA concluded, for instance, that it was still uncertain about whether the group of jobless people without basic qualifications actually benefited from the measures taken, as these measures did not provide for specific action to reach this target group. This issue was part of the original audit, and the fact that the NCA audit team found no evidence of success three years later is relevant to keeping alive the debate on how to best help jobless people without qualifications to gain better prospects to find a job.

CONCLUSION: FIRST STEPS TOWARDS THE BEST OF BOTH WORLDS

During the last decades, a lot has been invested in the development of performance management and results-oriented budgeting. In parallel with this, in both government and national audit institutions, performance auditing has established itself firmly alongside regularity and financial auditing. With many achievements has come some criticism, in particular about what is seen as too much of a focus on objectives and indicators will arguably limit our understanding of reality. To return to ‘Yes Minister’ – we should beware of congratulating hospitals that perform excellently on some indicators when in fact there are no patients.

Recent initiatives by the Netherlands Court of Audit to bring greater responsiveness into performance audits illustrate the difficulties in finding the best of both worlds. In addition to the old challenge of how to efficiently investigate the effectiveness of policy programmes, new issues around the quality, independence, and added value of performance audits arise. Nevertheless, the initiatives also show that more open and more participatory ways of programming, performing, and following-up systematically on performance audits, can help to balance the advantages and the risks of policy objectives and performance indicators. Selecting topics that are relevant to stakeholders, asking fundamental questions about the ‘fit’ between policy programmes (plus their objectives and indicators) and society’s problems, and looking back on the benefits of audit conclusions and recommendations help to keep both the auditor and the policy makers alert and responsive to changing contexts and preferences.

This being said, performance auditing should not do away with its professional audit norms and criteria. Neither should auditors join the chorus of those who deny any benefits from policy objectives, performance targets, and indicators. True, the traditional professional basis of performance audit – that is, the careful application of norms and criteria that constitute the basis for performance judgements – will automatically lead to a top-down approach in which objectives, criteria and indicators constitute the frame of reference for making judgements on policy intentions, implementation, and results are important to judge success. Yet at the same time, and somewhat paradoxically, this basis allows for more responsiveness to be included in performance auditing.

One of the most important tasks of responsive performance auditing is to keep government well aware of the ‘taming trap’: the illusion that, with the articulation of policy solutions through objectives, indicators, and criteria, problems may appear ‘under control’, while important negative consequences can be overlooked. Building upon the main questions from

the results-oriented budgeting system in the Netherlands (Have we achieved what we intended? Have we done what we should have done in achieving it? Did it cost what we expected?), on which – in principle – information should be available on a regular basis, the main set of questions should therefore be:

- Are we (still) on the right track?
- Are our assumptions (still) valid?
- Does the set of policy objectives and performance indicators we focus on (still) represent what we must try to achieve?
- Which negative side-effects of the use of policy objectives and performance indicators can be observed?

For this author, performance auditing must contribute to policy-oriented learning. Important questions therefore are: to which learning can performance audit best contribute? To the improvement of existing policy programs, or to discussions that may lead to renewal or even abandonment of those programmes? (Compare Van der Knaap, 2006.)

For the critical performance auditor, the challenge is to facilitate the combination of analytical focus, with the ability to continuously and critically review assumptions and performance indicators. In a results-oriented policy system, we need ‘verifying’ performance audits that explicitly raise the validity of existing policy theories, targets and performance indicators. The paradox is that this can be done while, at the same time, capitalising on the advantages of objectives, indicators, and norms. By seeking information and insights from stakeholders (for example, implementation offices, recipients), responsive performance auditing may provide feedback and new information on the validity of existing insights with respect to policy objectives and performance indicators. But it would be wrong to imagine that these questions can only be asked by performance auditors or policy evaluators. On the contrary, government, as a key part of a representative democratic society, has a duty to be open and organise responsiveness itself. But auditors need to both (a) audit the responsiveness in government and policy programmes, and (b) incorporate responsiveness as a key feature of selecting audit topics and conducting the auditing process itself.

NOTES

1. I stress the intermediate nature of findings, as the Board of the Court of Audit always chooses its own – definition and wording of – audit findings and recommendations.
2. Accountability Day is the third Wednesday in May when the annual accounts – including performance data and the outcomes of evaluations – are presented to Parliament by the

Minister of Finance, together with the President of the NCA. There follows a debate on the results of government policy programmes and spending which may lead to adjustments, etc.

Further information is at: <http://www.minfin.nl/dsresource?objectid=80394&type=org>

REFERENCES

- Abma, T.A. (1996), *Responsief evalueren – Discoursen, controversen en allianties in het postmoderne* ('Responsive Evaluation – Discussions, Controversies and Alliances in the Postmodern'), Delft: Eburon.
- Abma, T.A. (2005), 'Responsive evaluation: Its meaning and special contribution to health promotion', *Evaluation and Program Planning*, **28**, 279–89.
- Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale, and Burt Perrin (eds) (2007), *Making Accountability Work – Dilemmas for Evaluation and for Audit*, New Brunswick: Transaction Publishers.
- Bovens, Mark (2005), 'Public accountability', in E. Ferlie, L. Lynne and C. Pollitt (eds), *The Oxford Handbook of Public Management*, Oxford: Oxford University Press, pp. 182–208.
- Bovens, Mark and Thomas Schillemans (2009), *Handboek publieke verantwoording*, Utrecht: Lemma.
- Court of Audit (2005), *Essentials of Good Government*, The Hague.
- Dryzek, John S. (1993), 'Policy analysis and planning: From science to argument', in Frank Fisher and John Forester, *The Argumentative Turn in Policy Analysis and Planning*, London: UCL Press Limited, pp. 213–32.
- Fisher, Frank and John Forester (1993), *The Argumentative Turn in Policy Analysis and Planning*, London: UCL Press Limited.
- Frissen, P.H.A. (1996), *De virtuele staat*, Schoonhoven: Academic Service.
- Frissen, P.H.A. (2007), *De Staat van Verschil - Een kritiek van de gelijkheid*, Amsterdam: Van Gennep.
- Guba, Egon G. and Yvonna S. Lincoln (1989), *Fourth Generation Evaluation*, Newbury Park: Sage.
- In't Veld, R.J. (1989), *De verguisde staat*, The Hague: VUGA.
- INTOSAI (2010), *Implementation Guidelines for Performance Auditing - Standards and Guidelines for Performance Auditing based on INTOSAI's Auditing Standards and Practical Experience*, INTOSAI.
- Ling, Tom (2007), 'New wine in old bottles?', in Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (eds), *Making Accountability Work – Dilemmas for Evaluation and for Audit*, New Brunswick: Transaction Publishers.
- Majone, Giandomanico (1989), *Evidence, Argument, and Persuasion in the Policy Process*, New Haven: Yale University Press.
- Ministry of Finance (2004), *Evaluation of VBTB*, The Hague: Ministry of Finance.
- Pawson, Ray and Nick Tilley (1997), *Realistic Evaluation*, London: Sage Publications.
- Schwandt, T.A. (2001), *Evaluation Practice Reconsidered*, New York: Peter Lang.
- Schwandt, T.A. (2002), 'Democratization, globalization and performativity: Implications for learning in and from evaluation', Keynote address given at the 5th biennial meeting of the European Evaluation Society, 12 October 2002, Seville.

- Scriven, M. (1991), *Evaluation Thesaurus*, Newbury Park: Sage Publications.
- Stake, R. (2004), *Standards-Based and Responsive Evaluation*, Thousand Oaks: Sage Publications.
- Van der Knaap, Peter (1995), 'Policy evaluation and learning: Feedback, enlightenment or argumentation?', *Evaluation*, **1** (2), 189–216.
- Van der Knaap, Peter (2006), 'Responsive evaluation and performance management – Overcoming the downsides of policy objectives and performance indicators', *Evaluation*, **12** (3), 278–93.
- Van der Knaap, Peter (2007), 'Results-oriented budgeting and evaluation: Accountable for learning', in Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (eds), *Making Accountability Work – Dilemmas for Evaluation and for Audit*, New Brunswick: Transaction Publishers.
- WRR (Wetenschappelijke Raad voor het Regeringsbeleid) [Scientific Council for Government Policy] (2004), *Bewijzen van goede dienstverlening* [Proofs of Good Service Provisions], Amsterdam: Amsterdam University Press.
- WRR (Wetenschappelijke Raad voor het Regeringsbeleid) [Scientific Council for Government Policy] (2007), *Lerende overheid - Een pleidooi voor probleemgerichte politiek* [The Learning Government], Amsterdam: Amsterdam University Press.

PART III

Conclusions

15. Conclusions: performance audit – an effective force in difficult times?

Jeremy Lonsdale, Tom Ling and Peter Wilkins

INTRODUCTION

In the conclusion to ‘Performance or Compliance: Performance audit and public management in five countries’, Christopher Pollitt comments that in the 1990s performance audit ‘arrived’, becoming ‘a known quantity, a familiar presence in the process of governance.’ At the same time, he suggested that SAIs faced strategic choices in respect of: the core practices and criteria they used; the topics they selected for their studies; and their definition of, and relationships with, their prime and secondary audiences. Pollitt added that:

In different ways and at different speeds, each SAI has, through a mixture of pressure and choice, begun to emerge into the wider world. This emergence has involved dialogues with executives and legislatures, the adoption of more proactive policies towards the media, and gradually increasing cooperation with other SAIs. It is a trajectory which, sooner or later, leads to new comparisons being made and demands being put . . . (Pollitt et al., 1999, 219)

More than a decade later we have considerably more evidence on the trajectories that performance auditors have followed, and can take a longer perspective on the pressures and choices they have faced. We can also look beyond the world of SAIs and see more generally how performance audit has developed. We can do so because, as Furubo notes, although there is not an extensive literature on performance audit, there is now far more material available than when Pollitt made his observations. A comprehensive list of articles, books and papers on performance audit or related topics maintained by one of the contributors to this book (Vital Put) identified 257 items published between 1979 and 2008. Of these, 84 (33 per cent) appeared before 2000, with the majority published in the eight years from 2000. This book has added further to the evidence base,

with many chapters synthesising information from multiple sources and others providing new empirical evidence. This final chapter draws on this evidence to conclude on:

- what we have found;
- what contribution performance audit is making to accountability in democratic government;
- what the evidence says about the nature of performance audit; and
- where we consider performance audit and where those who create, use and are affected by the information should direct their attention in the future.

WHAT HAVE WE FOUND?

Overall, there is no doubt that performance audit has continued to develop over the last decade. It is now a stronger discipline in many countries, able to draw on wide statutory powers, and with a track record of impact. As we have seen throughout the book, performance audits examine a wide range of important, complex and pressing subjects in areas as diverse as, for example, health, education, defence, infrastructure, regeneration, transport, and the environment. Auditors have adapted their work to focus on new and emerging subject areas: privatisation, the use of private finance, the creation of public-private partnerships and the use of regulatory mechanisms. Most recently, they have started to report on the responses made by governments to the economic downturn, such as the rescue measures for the banking sector, fiscal stimuli and the support for different industries.

Performance audit has drawn on other fields of enquiry to become a more rigorous activity and has developed quality assurance arrangements which are at least the equal of those found in many other forms of evaluative activity (Lonsdale and Mayne, 2005). Auditors are now more alert than they were to the risks and downsides of its practices and perspectives. Thus, the performance audit community has come a long way, in our opinion, in understanding the potential and the challenges of its role, and is now more diverse, comprising a richer set of skills and professional perspectives than in the past.

We can look at this in a little more depth. Firstly, there has been a growth in the scale and intensity of performance audit work. It is now undertaken by more SAIs than ever and on a larger scale. The GAO, for example, undertakes around 1,000 performance audits a year, whilst during the mid-2000s the NAO increased its output from 45 to 60 major

reports a year. The other SAIs in our study continue to pursue significant programmes of work. In Australia, audit bodies carry out performance audits in most of the states and territories, as well as nationally, and in Canada, in each of the provinces. These are substantial projects; many of the audits make use of multiple data sources and more intrusive and complex evidence-collection methods than in the past.

One American overview (Burke and Haynes, 2008) suggested that whilst performance audit was 'less ubiquitous [than financial audit, it is] a broad and growing professional activity, as many governments feel the need to submit their physical operations to close scrutiny based on the values of efficiency and effectiveness.' Funkhouser's chapter charts the expansion of performance audit in cities in the United States, arguing that the growth is likely to continue as the work's main preoccupations – accountability and improvement – remain central to debates about the use of public resources. Weets highlights the establishment of a new performance audit regime in the Dutch city of Rotterdam. Elsewhere, we have seen the creation of audit bodies in Scotland and Wales, under devolution arrangements, which has increased the amount of performance audit undertaken in the UK and taken it in different directions. Many SAIs have started to carry out performance audit where there was none before; for example, in many developing nations. Some of the growth has been stimulated by projects funded, for example, by the World Bank, European Commission, the Asian Development Bank, or led by SAIs with well-developed performance audit regimes. There is, as Furubo noted in Chapter 2, no end of powerful advocates for performance audit.

Such developments might be seen as part of Michael Power's 'audit explosion', with the associated negative connotations, but there is evidence that there is an appetite for the work – from parliaments, governments, the media and community groups. One review identified that governments had found new roles for audit to play in a number of countries; for example, in Australia, New Zealand and the United Kingdom (Wilkins and Lonsdale, 2007). Dunleavy et al. (2009) found substantial press coverage for NAO reports between 2004 and 2009, whilst individual pieces of work can stimulate considerable interest; for example, the Office of the Auditor General's report on the Canadian federal government's management of a contract given to a communications agency, which led to intense public debate and political upheaval in Canada (Fraser, 2007).

It is unclear, however, how much performance audit is required, or what factors may limit its further growth. SAIs cannot take attention for granted. Nichol (2007) found limited interest in the Australian House of Representatives committees in ANAO performance audits, and Crerar (2007) reported greater familiarity with inspection and complaints

handling than regulation and audit amongst the public in Scotland. There are signs that some SAIs are reducing the amount they spend on individual studies and operating to reduced budgets. The moves in many countries to reduce significantly the size of their public sectors may stem the expansion of performance audit, forcing SAIs to reduce their own costs, cutting the size of their audit fields, and prompting discussion about whether performance audits are the most cost-effective ways of assisting government to improve. The increase over the last ten years has coincided with a period of economic growth and the expansion of public sectors. As we enter a period of restraint and retrenchment it will be interesting to see how the performance audit function responds.

Secondly, we have seen the continuing development of very distinctive approaches in different countries, driven in part by the needs and requests of legislatures and governments, the skills of SAI staff and external experts and the differing philosophies of the audit bodies. Notwithstanding the efforts to develop and popularise international standards, covered by Wilkins and Boyle, performance audit has continued to emerge in distinct forms and with different priorities, and although there will be continuing opportunities to learn across SAIs, there is no reason to expect that this will lead to a greater convergence in their approaches. Put, quoted in Boukaert and Halligan, 2008 has argued that, for example, the Dutch and the UK approaches to performance audit are almost exactly opposite strategies. In the UK, management issues are priorities, whereas policy evaluation is secondary. In the Netherlands, he argues the opposite is the case. As Put notes in Chapter 4, the Dutch SAI focuses on how ministers should make policy and give accounts to Parliament, and pays less attention to how public managers administer. Van der Knaap's chapter on responsive performance audit also illustrates a distinctive departure in The Netherlands. The 'reality checks' carried out in 2009 and 2010 have sought to assess the continuing relevance of policy measures, starting with the societal problem the government intervention is seeking to tackle.

The weight that SAIs attach to the influence of different stakeholders also pushes them into distinctive forms of performance audit. Put and Turksema's chapter on study selection points to the different impacts of institutional, organisational and operational influences which are driving SAIs in different directions, encouraging them to prioritise particular aspects of their work; for example, the need to demonstrate financial savings; or driving them to focus on particular themes as a result of their own forms of risk assessment. Similarly, Wilkins and Boyle in their examination of standards and quality conclude that the different standards in operation raise the potential for very different futures for performance

auditing. It therefore seems likely that local needs and interpretations will remain highly influential.

Thirdly, we have seen a broadening of methods and approaches. Burke and Haynes (2008) suggest that 'Ultimately, the performance auditing professional is as much an artist as she or he is a scientist, working from a broad mandate with an abundant toolbox of techniques and broad discretion in approach and focus.' Different approaches have been fuelled by separate recruitment practices, different views on what constitutes appropriate skills, and varied training regimes. Lonsdale (2000) commented that performance auditors in four SAIs in the Netherlands, Sweden, Finland and the UK had been at pains to highlight that they had given greater attention to how they applied methods in the late 1990s than in the past. Astill and Page (2006) also concluded that there was both pressure from the top of the UK and French SAIs to develop their methodologies, as well as effects from below 'due to auditors pushing out the boundaries'. Lonsdale goes further in this volume to identify a series of factors that shape the choices and constraints on auditors, and also highlights some of the innovations that have taken place within the work of the UK NAO. Factors influencing the choice of methods include the practical constraints of time, cost and evidence expectations, but also multiple influences at team, project, organisational and environmental levels which have stimulated the use of new methods. Scharaschkin also underlines why the appropriate use of methods to generate robust evidence is essential, likening the process of drawing conclusions for audit to that required for making a legal judgment, and highlighting the need for auditors to be able to withstand often heavy challenge to their findings from those being audited.

Frans Leeuw's chapter, however, expresses doubts about the degree of sophistication in the methodological quality of some performance audit work, picking up a theme referred to in the past (Roberts and Pollitt, 1994, Schwartz, 1999). He highlights a lack of theory-driven performance audit and a failure to address the attribution issue, also arguing that the theory underlying performance audits is seldom articulated. He warns that this raises questions about the validity of statements arising from performance audits. His concerns are aligned with those raised by Dubnick (2005) (mentioned in this volume by Funkhouser) that there is nothing in the literature to provide a logical link between account giving and performance. Yet performance audit is not alone in facing challenges in establishing cause and effect in complex systems, and the progressive developments in the methods used indicate that many SAIs are facing up to these challenges. How far they will be able to go remains to be seen, however, given that, as Lonsdale comments, there is growing pressure on SAIs to deliver to tighter timetables.

A fourth theme has been the pursuit of higher standards. In many ways we can see parallels with the debate in the world of evaluation, with SAIs seeking to codify, standardise and communicate the expectations to which they work. As Wilkins and Boyle discuss, INTOSAI has developed performance audit standards and individual audit offices have either developed their own or chosen to adhere to established ones. An unresolved issue is whether a greater consistency of standards and practices will emerge in coming years or whether the diversity of approaches will continue. In the United States, Burke and Haynes (2008) have seen 'a movement in recent decades to insert a common framework through legislative and bureaucratic action', with the GAO's generally accepted government auditing standards (GAGAS), first released in 1972, becoming the norm for national auditing, and state and local governments, as well as higher education institutions and non-profit agencies, receive federal assistance being required to use GAGAS. There is certainly a desire for consistency and a need, given the increased expectations of transparency under freedom of information legislation (and even court proceedings as in Canada (Fraser, 2007)), to be able to show the standards that have been applied.

The pursuit of quality in audit work remains an integral and continuously challenging part of the role of SAIs. Study design has been adapted to cope with new permutations of policy, new trends and even fashions in public management, new causal processes in society, new political demands for assurance and robust performance information, and the availability of new technologies for data collection and analysis. The recent rapid tempo of change in societal and government information systems is highly unlikely to slacken in the next decade, as the spread of 'digital era governance' institutions and solutions proceeds apace (Dunleavy et al., 2006). The onus will be on all SAIs to remain competitive and authoritative in their performance audit work with many other sources of information about performance in government and society. Some of these sources will demonstrably not have the same credibility and rigour as others, but will nevertheless help to shape public discourse, perhaps because they will be available more quickly or in more easily digestible forms. This is likely to drive SAIs to emphasise the quality of their work and the standards to which they operate, given the need to demonstrate their value.

Another response could be for SAIs to be less concerned with being the authoritative source of evidence about performance, and more concerned with establishing themselves as the source of authoritative and independent judgement based on this evidence; their role would be to make sense of the burgeoning body of performance data. At the same time, we would advise against SAIs becoming over-preoccupied with codification and

standardisation. While consistently high standards are obviously desirable, it is less clear that more precisely codifying standards and practices is necessarily the best way to achieve this. An absence of explicit standards has not necessarily prevented the development of performance audit. For example, it was only in 2008 that the NAO in the UK introduced its own VFM standards, and in 2009, Audit Scotland. Nevertheless, both have well established performance audit work.

Fifthly, we have seen increased interaction between SAIs and the outside world in developing their performance audit. Lonsdale (2008) has argued that the history of VFM work at the NAO can to a great extent be seen as the development of a range of external contacts. In particular, relationships have become:

wider than in the past (beyond the traditional connections with Parliament and the audit profession); deeper than in the past (in terms of their significance for the work, eg academic quality assurance and partnership working with external providers); more complex in their demands on the NAO (being a mix of statutory, professional and voluntary arrangements); and more likely to create tensions (eg some encourage conservatism whilst others encourage innovation).

Performance auditors have increasingly sought to explain the conduct of their work to others. In 2008, the OAG in Canada prepared a detailed document for the staff of audited departments entitled 'What to Expect: An Auditee's Guide to the Performance Audit Process', designed to:

encourage productive and respectful relations between entities and audit staff. The tone of the desired relationship is one of mutual respect and the document sets out the roles and responsibilities on both sides. The objectives for the auditor are to:

- make an ongoing and consistent effort to understand the context in which government departments and agencies do their work,
- promote open two-way communications, and
- act in a professional and objective manner.

Wilkins and Boyle highlight another aspect of this in the use of external reviewers – from other audit bodies or from academic institutions – to assess the quality of performance audits or the performance audit function. The performance audit reports of the Office of the Comptroller and Auditor General of Ireland, for example, were reviewed by external experts, who drew up an assessment system based, amongst other things, on the MEANS grid used by the European Commission and the model for rating departmental performance reports used in Canada. This has helped to ensure that wider perspectives on quality and rigour are reflected in the expectations set for SAIs, drawing them into wider methodological

debates. The influences from beyond the performance audit community can be seen in other ways. Lonsdale's chapter on methods shows how the NAO have been influenced by evaluators, management consultants and academics in integrating new methods into their work.

Furubo's chapter though reminds us that there is still a risk that performance audit might remain an insular discipline, not subject to the same self-examination, and cross-pollination of ideas as, for example, evaluation, political science or public management. Whilst there is an extensive network of audit institution conferences and workshops, they tend to be attended almost entirely by auditors and be technical in content, rather than reflective. The growth of many other forms of evaluative information highlights the need for performance auditors to communicate their messages clearly, stressing what it is that makes their contribution unique.

A sixth and final theme has been the need to tackle increased complexity in the subject matter addressed. Van der Knaap comments that if performance audit is to contribute to effective government, 'it must pay close attention to the complex and dynamic nature of public policy'. Such complexity includes different centres of power, varied populations with different cultural heritages, technological developments and constantly changing administrative arrangements.

Performance auditors see their work as being subject to such changes in their environment. In the words of INTOSAI:

Performance auditors can be faced with considerable variety and ambiguity in their work. They require skills in analyzing activities and management practices. They can be faced with the need to become familiar with a wide range of organizational contexts and subject matters. They need the ability to write logically and thoroughly on complex issues. (INTOSAI, 2004)

Similarly, an NAO guide advised staff:

There are many challenges ahead. Government itself is now far more focused on delivering value for money. The political debate about public services has raised the profile of our VFM work, which reinforces the need for our reports to be precisely worded and our conclusions independent and soundly based. Judgements about value for money are becoming more complex, with different public service delivery models and a greater desire on the part of government to change individual behaviour, rather than simply deliver services. The expectations of our stakeholders are more sophisticated and there are greater demands that our work will lead to lasting improvement. And there is more information available than ever on government performance; at times, we are competing for attention with others who know a lot about a topic. In these circumstances, expectations of our VFM work are growing and we must meet these demands. (National Audit Office, 2008)

Performance audit has also extended to parts of government activity, often specialist and complex, where it was previously not seen; for example, the role of private sector providers and their contracts with the state. We have also seen the NAO in the UK examine the role and performance of utility regulators (sectors of the economy which interestingly were not subject to state audit oversight when under government control), the work of the voluntary sector where it is working for government, and environmental audit. Each has required application of new skills and expertise.

We also see some audit offices now tackling the work of governments to deal with the economic crisis. In May 2009, for example, the Algemene Rekenkamer published its first report (Algemene Rekenkamer, 2009) on the main interventions and arrangements made by the Ministry of Finance in response to the credit crisis. In May 2010, it reported on the further development of these interventions and arrangements (Algemene Rekenkamer, 2010). The National Audit Office in the UK has reported on the nationalisation of the bank, Northern Rock, wider support for the banking sector and the effect of the credit crisis on infrastructure development (National Audit Office, 2010). And the Swedish NAO reported in 2009 on 'State Guarantees in the Financial Crisis', concluding that the government had not provided comprehensive and transparent status reports to parliament setting out the risks for the state and taxpayers (Riksrevisionen, 2009).

It has not always been easy. The difficulties and risks – both methodological and political – with such work can be seen, for example, in the criticism of the work of the audit offices in Australia in scrutinising public–private partnerships, where there were perceptions that they 'may have compromised their independence through their attempts to foster "good management practice" in the public sector' (English, 2007). A UK NAO report on the health service's National Programme for IT – a highly complex and controversial project at an early stage of development – also led to criticism of the auditors, with accusations made in the media that the report's findings had been watered down under pressure (Dunleavy et al., 2009).

The challenges of this greater complexity and SAIs' responses to it have led some to question whether performance audit and the existing institutional arrangements are being stretched too far (Ling, 2007). As it becomes clear that outcomes are in part shaped by complex systems, holding people and organisations to account for outcomes over which they have no meaningful control can lead to a spurious accountability. Performance audits that do not address whether worthwhile outcomes are being achieved, and restrict themselves to whether the public sector is doing a good job administering the policy of the government of the day, may disappoint their readers,

especially specialists, who may be looking for coverage of issues such as adaptability, responsiveness and effective learning. However, even if elected assemblies and the public they represent want their SAIs to move into more complex territory, it is not clear that SAIs are always well placed to investigate these questions. There is, therefore, a dilemma facing SAIs posed by the growing complexity of public services: to ignore it risks falsely attributing outcomes to actions, but to engage with it can take SAIs into questions such as ‘Did learning take place?’ or ‘Does the organisation have the capacity to perform into the future?’, where the evidence needed to support the audit judgement is hard to find and have agreed by stakeholders.

WHAT CONTRIBUTION DOES PERFORMANCE AUDIT MAKE?

We now move on to consider whether the changes in performance audit have made a difference to the quality of government. In particular, we consider:

- whether there is evidence that it is having an impact on the performance of government; and
- whether it is a force for good in democracy.

Both of these questions require careful consideration. The first because there is no ‘counterfactual’ allowing us to compare the performance of a particular government with and without performance auditing. The second because, in addition to the problem of the counterfactual, the meaning of ‘democracy’ is itself contested.

Performance Audit and its Impact on Government Performance

Several chapters in this book provide contrasting views on the impact of performance audit. These differing views are in themselves important. Like other forms of evaluative activity, it seems performance audit in one context can produce demonstrable benefits, but in another it produces limited, or even negative, outcomes. This perhaps explains the variety of views expressed here and in the wider literature; some contexts appear to be ‘audit ready’ and able to absorb and benefit the lessons from SAIs. Other situations, perhaps because they are in rapid flux or suffer from an inability to learn and adapt, may be unable to benefit from even the most perfectly crafted performance audit. Part of the art of conducting the work is, therefore, to judge what will add real value.

It is clear that both what makes a good performance audit, and when the context is ready to benefit from it, are contested issues and ones on which debate will continue. We have sought to highlight several perspectives. Funkhouser's chapter, for example, is built around exploring the contrasting views of academics and practitioners on the value of performance audit. As he noted, where academics have commented on performance audit they have 'employed a range of arguments, from democratic theory to extensive empirical evidence, to persuade us that performance auditing's impact on public organizations is at best benign, and at worst malevolent'. In contrast, auditors that he quoted from across the United States, like those interviewed from the NAO (Lonsdale, 2000 and in Chapter 13 in this volume), all cited examples of where action has been generated by their work, saving money and improving services. These sources also commented on the well-intentioned approach of auditors and the desire to make a difference as they saw it. On the other hand, Weets' chapter noted mixed views amongst those audited as to the value of audit work. She highlighted occasions when auditees felt auditors acted too much according to the letter of the law, failed to take context sufficiently into account when presenting their findings, and where a lack of empathy with administrators harmed the technical quality of their work. It is clear from Weets' and Scharaschkin's findings that relations can be tense, even confrontational, and Lonsdale and Bechberger also noted that there can be scepticism from auditees at attempts to go beyond what are seen as the formal requirements of the audit work. All this evidence suggests there are communications challenges regarding their intentions, as well as a relationship-building role for auditors.

Van Loocke and Put analyse in detail what limited empirical data is available on the impact of performance audit. Their work suggests, perhaps not unexpectedly, a mixed picture, with impact often 'slow and subtle', at times quick and explicit, and sometimes, hardly noticeable. Much of the impact is instrumental, rather than conceptual; some of the impact can be slow to materialise and does so in a complex way. They conclude that there is still little understanding of why some pieces of work are more successful than others, but highlight factors that seem to be influential: the relationship between auditors and auditees; the timeliness and relevance of the results; the influence of third parties such as parliaments and the media; the evaluation culture in the audited body and the coincidental timing of audit and reform projects underway. And they also believe that when measuring impact, underestimation, rather than overestimation, is most likely.

Thus we conclude that impact is happening as a result of numerous performance audits, but not always or as often as would be desirable. We also

emphasise that there are actions that auditors need to take to maximise their impact: focusing attention on the quality of the auditor–auditee relationship, building ‘advocacy coalitions’ with others, selecting topics very carefully and putting effort into disseminating their results. Being aware of the risk of adverse impacts of audit, and mitigating them where they arise, is also important. Leeuw’s chapter highlights a variety of potential effects, and although Van Looke and Put argue that the relevant literature amounts to not much more than a form of risk analysis, there is no doubt that auditors need to be aware of the risks of, for example, discouraging innovation and encouraging gaming and resistance. If we were to advocate a priority for a future research agenda then gaining a better understanding of the circumstances in which performance audits do, and do not, have impact seems most important.

A belief that the work has impact should not, however, blind us to the possibility that the usefulness of performance audit in its current form may come to an end. We have clearly moved from a long period of incremental growth (and occasionally incremental contraction) of the public sector associated with new programmes and delivery mechanisms. The focus for scrutineers here has been on efficiency gains and recommending refinements, rather than major changes. It is conceivable that there will be a limited audience for this kind of report in the future. We have now entered a more turbulent time, characterised by cuts in programmes and retrenchment, which could conceivably require a different product or form of analysis. SAIs will need to guard against institutional inertia leading them to keep churning out reports of the same kind, and focusing on topics of little relevance to the contemporary issues facing parliaments and governments. The flexible nature of the product, the skills base and the governing legislation suggest they stand a good chance of being able to adapt, but it will require careful thought and a sound awareness of the changing environment.

Performance Audit and Democracy

If SAIs are to remain relevant they must have a specific and unique contribution to make to democratic government. In representative democracies we see the election of an assembly that holds the government of the day to account. Performance audit can and does contribute to this holding to account where reports are used in different ways as the basis of parliamentary hearings or debates. Audit offices are thus locked into democratic arrangements, albeit as secondary players, but with an important role to play. They directly improve the quality of information and argument available to elected representatives to hold government to account. At

the same time, they also improve the evidence and arguments available to interest groups, political parties, the media and citizens so that they can also select, inform and direct their representatives. Democratic debate is well served through relevant and reliable information, which becomes even more important when the focus of much of this public debate is less about ideological difference and more about managerial performance.

We can break the contribution of SAIs and performance audit to democracy down further in terms of the relevance to: accountability, trust, the availability of independent information, and the relationship between government and citizens. Accountability is an essential precondition for democracy. The role played by performance audit and evaluation in accountability arrangements has been explored in detail elsewhere (Bemelmans-Videc, Lonsdale and Perrin, 2007). They argue that:

Accountability can serve both as an end in itself, and as a means to a greater goal. A major purpose (some would say its prime function) is the legitimization of the exercise of authority, including the most appropriate use of public resources. The relationship is not simple however, and there are concerns that despite the increased resources and attention given to accountability there is often a sense that government or that parts of government (eg. quangos) are less accountable. In these circumstances, performance audit can be torn between competing visions of what makes for effective accountability, in particular, the balance between what has been seen as the 'culture of compliance and following the rules' and the more 'results-oriented approach that reinforces taking measured risks, learning and constant improvement'.

This issue has arguably become more contentious. For example, how can we hold governments to account for complex and at times hasty interventions such as the response to the economic crisis (for example, recruiting expensive financial advisers at short notice), where the alternative to the actions taken may well have been economic catastrophe? At a time of considerable uncertainty, and when many are looking for reassurance, those who want to 'tell it as it is' are not always welcomed.

This book has argued that the prime focus of performance audit is accountability. Generally, the work has been successful in this role by being based on independent and reliable information on the activities and results of government, which sets out how government initiatives have fared and by providing material that has been the basis for legislatures and their members to probe and question. SAIs have used their independent status to select topics for performance audit that have gone well beyond pure compliance and the following of rules, to a far wider range of concerns. They have reported on success and failure in policy implementation, exposed poor practices and performance, as well as the failure to learn from experience. Contrary to the perception of some, they have generally sought

to avoid being unnecessarily critical and, as van der Knaap and Lonsdale and Bechberger suggest, have been taking steps to take a more responsive approach. Overall, performance audit has played a significant role in ensuring that government is held to account for its use of public resources.

A second aspect of performance audit's relationship with democratic government relates to the issue of trust in what governments say. Various commentators have suggested that trust in society is in decline; Onora O'Neill has spoken of 'our present culture of suspicion' (O'Neill, 2002), and politicians, professionals of many kinds and other authority figures have seen their standing fall (Seldon, 2009). In these circumstances, those responsible for reporting on the performance of government and tasked with providing assurance on whether public funds have been spent as anticipated have an ambiguous role. As De Lancer Julnes (2006) has argued, 'The current emphasis on results for both accountability and performance measurement has been associated with the increased skepticism and discontent of the American public with how their tax dollars are being spent'.

The relationship between performance audit and trust is contentious. From one perspective, performance audit reports which regularly identify deficiencies and failures weaken trust in our public institutions. Alternatively, it can be argued that performance audit reports, in placing information in the public domain, have helped to increase trust in democratic government (or show where it is misplaced). Published audit reports provide information and analysis independent of government and the existence of the function gives general assurance that someone is keeping an eye on how public money is spent. Our view on trust depends on our view of public servants and whether we regard them as 'saints', who are best improved through guidance, or 'rascals', who need to be controlled. In truth, it probably takes an odd view of the world to think we can trust everyone without evidence. And we might argue that trust is not good 'in itself', but only has significance when it has been authentically won. Our trust in officials should be enhanced by knowing what they and their organisations are doing, and knowing that scrutiny of them has been conducted independently. Transparency and independent reporting allow this, and it is possible to undermine trust, but at the time contribute to transparent representative government, for example, by stopping governments overclaiming their achievements. Of course, limitations in public sector performance are not all due to the failings of individuals, and the public trust also relates to information on whether the systems that are in place to deliver government policies are appropriate and adapting to changing needs. By reporting on the reliability of systems and practices, as well as performance, and by making their analysis and reporting accessible

to the non-specialists, we would argue that SAIs are helping to enhance trust in government, even if it can be uncomfortable for those in positions of responsibility.

A third dimension concerns the independence of information that SAIs issue in their performance audit reports. Dahl (2000) has commented that 'Like freedom of expression, the availability of alternative and relatively independent sources of information is required by several of the basic democratic criteria.' He questioned how citizens can acquire the information they need in order to understand the issues if the government controls all the important sources of information, arguing that citizens must have access to alternative sources of information that are not under the control of the government or dominated by any other group or point of view. A key strength that SAIs have is their ability to secure access to information, and also generate their own from within the public sector. Their reports also cast an independent eye over government claims. In 2010, for example, the UK's NAO reported on reported VFM savings made by government departments under savings programmes. They reported that 38 per cent fairly represented sustainable savings, 44 per cent may represent savings but with some uncertainty, and 18 per cent did not represent, or significantly overstated savings (National Audit Office, 2010).

The task for SAIs is to guard against accusations that they are not independent in how they carry out their work. The methods used by SAIs require the information reported to be assessed for its robustness and appropriateness; for instance, the GAO has very strict requirements on the use of third-party data. The NAO has been criticised by its external reviewers and some MPs for making too much, possibly uncritical, use of government data. This has led to greater expectations of setting out in reports what validation has been done or providing an indication of the independent quality assurance checks to which the data is subject. It emphasises that SAIs need to avoid relying too much on data from others.

Finally, performance audit plays a role in opening up government to examination by citizens. Providing information about public sector performance has grown in the form of increased data in annual reports, league tables, and summaries of information in local papers and the internet. As one observer has put it, 'Governments now attempt to show their constituents what they are getting for their tax dollars, how efficiently and effectively their tax dollars are spent, and how expenditures benefit constituents' lives' (Callahan and Holzer, 1999). Mulgan too (Mulgan, 2006) has argued that:

the best modern governments open themselves up to doubt and learning . . . This does not imply limitless tolerance . . . Instead its defining feature is that

it tolerates many paths to truth, and tries to internalize within itself sufficient challenge and argument to protect it from its own beliefs, with officials and advisers who are willing to speak truth to power. . . . the true interests of the state lie in scrutiny with the eye of an enemy. This is the principle long applied to companies to ensure that their accounts are truthful. It has been applied to professions and public services, particularly where there are big asymmetries of knowledge between citizens and those they depend on.

A number of SAIs have argued that they take a citizens' perspective. The Swedish NAO, for example, states 'One important role . . . is to ensure democratic transparency, i.e. provide citizens with the opportunity to see how democratic decisions are made and implemented, how their tax money is used, and whether public administration follows directives, rules and regulations and achieves the objectives set for it' (Riksrevisionen, 2010).

Overall, performance audit provides a unique flow of a certain kind of evidence and argument. Through its efforts to provide dispassionate accounts based on evidence and argumentation it lends itself to a particular, more deliberative, view of democratic life, which recognises the practical importance of majoritarianism but privileges the importance of authentic deliberation as the basis for decision-making. Arguably, where it does it best is where it reports complex material in ways that non-specialist readers can understand.

THE NATURE AND MEANING OF PERFORMANCE AUDIT

Having considered the evidence on the development and impact of performance audit, we believe we are getting closer to explaining the nature and meaning of performance audit. Furubo gave us a working definition in Chapter 2 – 'Performance auditing is an evaluative activity which produces assessments regarding performance, or information about performance, of such a reliable degree and with such a freedom from investigatory and reporting constraints, that they can be used in the realm of accountability' – which we consider holds up. The contributions to this book have highlighted different aspects of performance audit and support the constituent parts of the definition. Different chapters have provided support for the view that it should be: evidence-based – auditors must be able to substantiate what they say; analytical – it is not just evidence gathering as auditors must be able to derive meaning from their examination; resource-focused – examining how public resources and money have been used, how choices have been made about the use of funds; evaluative

– a judgement of performance and not just a statement of fact; prescriptive – auditors should in general conclude and make recommendations; and purposeful – supporting accountability arrangements and seeking to stimulate improvement.

It is also clear from the chapters in this book that the primary focus of performance audit is on accountability, with learning and performance improvement as an important, but nevertheless secondary, objective. It is the accountability role that provides the status and standing for the audit reports, and the opportunities to contribute, that engages representatives from governments and parliaments in a very distinctive way, and which creates particular opportunities for learning, even where it may also cause tensions and resistance. Immediate lessons are offered directly to the auditee from scrutiny, debate and exposure of data, but performance audit also has an opportunity to contribute more widely to decision-making and political life because of the flow of a particular type of information into official and professional communities and the media. Yet we would stress that the learning from performance audit takes place because it has been generated out of the accountability function, and is not secured in its own right. If well managed, the secondary learning function will amplify and strengthen the primary accountability role, but it is not the end in itself. As Furubo argues, if the primary purpose of an examination is learning, no one would ask for a performance audit.

Thus, we have a theory of change that links performance audit to better outcomes and better accountability, as follows. Independent, well-researched audits are produced which focus on appropriate issues. A legislature is made aware of problems and issues within government in a way that it would not have been otherwise, for example from internal investigations. Responsible bodies are overtly held to account and commit publicly to change behaviour or consider how such change could occur. In many cases, they do make these behaviour changes, which they would not have done otherwise. From this experience, wider learning and cultural change occurs within these bodies, as well as others that become aware of the performance audit and consequential changes.

This combination of contributing to both accountability and learning makes performance audit unique. It means that SAIs, although just one player in democratic accountability arrangements, or one voice amongst many advocating organisational improvement, can make a contribution which no other body can. In particular, we consider:

- They have very detailed insider knowledge of a kind that no other independent body has, both widely across government and in depth inside individual organisations. They also consider both

performance and the use of resources. These combinations make them authoritative about the effective and proper use of resources across the expanding range of models of public service delivery, an aspect of inquiry that is often weak or avoided altogether by other evaluators.

- They have a comprehensive understanding of the worlds in which they operate. Their status, position, powers of access to information and the longevity of their oversight, all mean they have unique insights into many different parts of government, based on accumulated experience.
- They are technical specialists but, at the same time, they also aspire to make their performance audit work easily understood by non-expert audiences, a feature in contrast to much research and evaluation, which can be characterised as professionals talking to each other. They are thus well placed to draw back the veil of confusion, complexity and obfuscation that so often gets in the way of a proper understanding of government.
- They can offer a cross-government perspective. Through their wider audit function they typically obtain information on all public sector organisations at least annually and monitor closely changes to legislation and policy. In theory, they can range over any combination of issues and organisations. This is in contrast to many research groups and think tanks which are more narrowly focused.
- They are well versed in analysing what we might call policy disasters or failures, and can often do so faster than academics, and in a more independent manner than in-house scrutiny. This is particularly important when there is a reluctance to own up to problems in government, or conversely, when in the media there is a tendency to dramatise every weakness. Their privileged access to documents and people, their objectivity, and their ability to publish despite opposition, all help to make them well versed in reporting on difficult subjects.
- They have flexible legislation, which should enable them to continue to adapt to changing forms of governance with broadly the same approaches, and be less adversely affected by changing fashions in methodological approaches or upheavals in public administration.

At the same time, however, we consider there are limits and constraints to the contribution that SAIs can make through their performance audit, and there are areas where SAIs are less likely to be successful. For example:

- They are likely to struggle in subject areas where purposes are ill-defined, where the evidence base is weak, making it hard or impossible to form a judgement, or where they are expected to answer general policy questions. They are also likely to be less effective where the central issue is major misconduct or corruption, where fraud investigators, the police or other judicial functions are better placed to act and report. And where the questions are too political, performance auditors may seek to avoid central aspects of the subject or will ignore the realities because they are too sensitive.
- They are likely to be less successful where their work runs the risk of being too technocratic and specialist. They may frame a question narrowly without taking account of a broader context, and thereby not identify relevant causal factors. For example, assessing the hastily developed stimulus interventions in response to the global financial crisis as if they were more conventionally developed programs would miss the higher level objective of maintaining community confidence in the economy, as well as the exceptional time constraints involved. There is also a risk of taking a narrow professional perspective, for example, in the current debate about the nature of the standards that should apply.

As well as these constraints, there are a number of threats to performance audit. The first is one of competition and of their work being lost amongst the many other voices commenting. At times, performance auditors can compete for attention in a context often saturated with information, which raises the possibility of being just part of the noise. This is a wider problem within representative democracy that noise drowns out the information needed for wise decisions. SAIs cannot solve this problem but must protect themselves by having a sharp focus on what they do that is distinctive.

Performance auditors must also be mindful of the underlying factors limiting public sector performance and the reporting of its performance. There is an interdependent relationship between performance auditing and performance measurement. They should consider the extent to which performance audits are shaped by the quality (or lack of) of performance measurement by agencies and the extent to which it leads to improved performance measurement by agencies. There should be a vision of what a well-running public sector should look like so that there could over time be a reduced need for performance auditing.

And they should avoid being too pragmatic, and too flexible and accommodating, trying to work in too many fields and moving too far from their core areas. They need to be able to recognise the limitations

of performance audit, and understand when time, cost and evidence constraints make a subject impossible.

AGENDA FOR THE FUTURE

In Chapter 1 we suggested that recent events had highlighted three key issues of great importance to government. These were: concerns for the value to be achieved from public expenditure; the mechanisms for accountability for public sector performance; and the credibility and trustworthiness of government and the information provided by it. We consider this book has demonstrated that performance audit has a major role to play in addressing and offering solutions to all three. Nevertheless, we want to end this chapter and this book by highlighting what we see as the agenda for the future for performance audit, and for those who create, use and are affected by the information, which will allow SAIs to develop further their democratic accountability role.

1. *Define what performance audit is and stick to it.* We think it is essential that SAIs focus on their core activities – and ‘be an SAI’. This will allow them to provide a unique source of evidence and judgement. SAIs should not aspire to be a social research body or to offer what could be termed ‘informed journalism’. There are more than enough think tanks in existence and enough observers commenting on public-sector performance, so SAIs should not get distracted by seeking to compete with them or operate in their sphere of influence. This requires them at times to be clear that there will be occasions when they are not best placed to examine a topic. They must protect their independence to avoid ‘regulatory capture’ and accusations of bias or partiality to particular interests and policies.

To this end, SAIs should be very clear what they consider performance audit is, be ever mindful of its role in democratic accountability, and should focus on executing it in an exemplary manner. They should also understand in what contexts they are most likely to be effective and able to deliver best value for money. Drawing on our definition of performance audit, they should be clear that the prime purpose of their work is to assist with accountability, with learning as an important secondary objective. They should champion the values of transparency and accountability as key elements of democratic government. They should also be focused on assessing public sector performance with a strong emphasis on resource use. A concern with ‘money’ as much as ‘value’ is in part what will make SAIs unique, especially when

there is so much concern on this issue, and when much other evaluative activity is often focused on what works, and not so much on how much it costs. But SAIs should also recognise the need for ongoing innovation and the need at times to be flexible and adaptable.

2. *Improve the theory and practice of performance audit.* Performance audit is a practical discipline, designed to assist in improving the operation of government. Nevertheless, it is important that it acknowledges and draws on emerging intellectual and methodological developments, and several chapters recommend further scholarly research to better understand aspects of performance audit. Leeuw suggests that there is a 'lack of intellectual depth' in many performance audits, and part of the developments recommended should, for example, be greater attention paid to cognitive-behavioural mechanisms that influence decision making. In line with Scharaschkin's analysis, there is also scope for further thought to be given to the structuring of argumentation within performance audits.
3. *Linked to this, develop the ability to assess and measure the impact and influence of performance audit work.* Every public body needs to be able to explain and evidence the benefits secured from its work. As we have seen, to date, much of the measurement around performance audit has focused on numbers of reports or recommendations 'accepted'. Too little attention has been given to outcomes or to a wider set of considerations associated with the influence of the work. The model developed by Morin and used in this volume by Weets offers a comprehensive way of doing it. It focuses on a wide range of indicators, including perceptions and environmental conditions. Based on a clear understanding of the impact and influence of performance audits, further work is needed to help guide parliamentarians and the wider community on appropriate levels of resource for performance audit at a particular time.
4. *Linked to this, tackle the downsides associated with the work.* To play a valuable role, performance audit must address a number of factors that help to reduce its effectiveness of the type discussed in detail by Leeuw, Van Loocke and Put and Funkhouser. SAIs must gain a better understanding of the impact and effect – good and bad – and the costs of compliance, and eliminate overlaps and duplication. Particular attention needs to be paid to the role of, and evidence base for, performance audit recommendations to ensure that they do not become policy statements or take inadequate account of alternative approaches to addressing issues raised by the findings.
5. *Improve the role of performance audit in respect of learning.* Our definition of performance audit places learning from the work as a

secondary occupation but, nevertheless, highlights its importance. We are clear that SAIs are uniquely placed to assist with learning because of their status and standing, their cross-government position, the depth of their knowledge of government bodies and the strength of their access to data and their analytical skills and judgement. This will require more explicit attention to be given to understanding the evidence base behind the 'good practice' material generated, so as to avoid merely publicising interesting examples they came across 'on their travels', for which there is insufficient validation of merit. They must also guard against 'mission creep' and moving into areas where other agencies have more to contribute to good democratic government. SAIs should also seek to maximise the learning that can be derived from the accountability processes they support, trying to help them be constructive in tone and forward looking, as well as demanding.

6. *Continue to increase their responsiveness.* This may require finding new ways of working to reduce the time taken, new means of reporting in line with developing knowledge of what kind of information works best with decision makers (Pollitt, 2006) and new ways of approaching their core roles. This requires a self-critical mind-set which acknowledges the changing nature of democratic scrutiny, the pressures on decision makers to react quickly and the availability of information in the public domain on an unprecedented scale.
7. *Exploit their position in the interests of deliberation and democratic decision making by striking a balance between supporting accountability processes and generating learning.* This can be done by:
 - (a) focusing on problems and issues stretching beyond individual organisations and programmes – in order to pick up, for example, the collaborative and cross-government issues;
 - (b) contributing to well-informed debate – for example, by producing work that helps to improve citizens' understanding of where taxes are spent and how different levels of government fit together, and by explaining systems and processes that currently confuse and alienate;
 - (c) supporting democratic citizenship – through the generation of public discourse around public performance by interpreting data, encouraging objective analysis, disseminating findings widely and making performance information more intelligible;
 - (d) asking inconvenient questions to keep government responsive to citizens – in a world of news management the ability to bring difficult issues to public attention is a crucial, if at times uncomfortable, one to be used wisely;

- (e) championing transparency – for example, by producing information that is value added, pushing for maximum information to be published, and writing reports in language that can be understood;
- (f) contributing to the creation of an environment for proper debate about performance – for example, maximising the benefit from developments such as the development of accruals accounting and performance information, as well as encouraging a mature debate about the causes of failure in government.

Ultimately, the role of performance audit should be to inform political, policy and management debates. But it should also be to give citizens grounds for having confidence in government, and the information available to them, showing them whether they are getting services that provide value for money, and are efficient and responsive. It must also be about explaining why, where appropriate, such confidence is unfounded. SAIs are well placed to take their unique but delimited role forward.

REFERENCES

- Algemene Rekenkamer (2009), *The Financial Crisis 2008–2009*, The Hague: Algemene Rekenkamer.
- Algemene Rekenkamer (2010), *Financial Crisis: Interventions and Follow-Up: Fourth Quarter of 2009 and First Quarter of 2010*, The Hague: Algemene Rekenkamer.
- Astill, Stuart and Ed Page (2006), 'Aims, Instruments and Interaction in VFM Audit: A Cross National Study', Paper for the National Audit Office, London.
- Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale and Burt Perrin (eds) (2007), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Brunswick, NJ: Transaction Publishers.
- Burke, Brendan and Wendy Haynes (2008), 'Performance audits', in Jack Rabin and T. Aaron Wachhaus (eds), *Encyclopaedia of Public Administration and Public Policy*, Second edition.
- Callahan, K. and M. Holzer (1999), 'Results-oriented government: Citizen involvement in performance measurement', in A. Halachmi (ed.), *Performance and Quality Measurement in Government. Issues and Experiences*, Burke, VA: Chatelaine Press, pp. 51–64.
- Crerar, Lorne (2007), 'The Report of the Independent Review of Regulation, Audit, Inspection and Complaints Handling of Public Services in Scotland', Edinburgh: Scottish Government, available at: www.scotland.gov.uk/Publications/2007/September (accessed 13 May 2010).
- Dahl, Robert (2000), *On Democracy*, Yale: Yale University Press.
- De Lancer Julnes, P. (2006), 'Performance measurement: An effective tool for government accountability? The debate goes on', *Evaluation*, 12 (2), 219–35.

- Dubnick, Melvin (2005), 'Accountability and the promise of performance: In search of mechanisms', *Public Performance and Management Review*, **28** (3), 376–417.
- Dunleavy, Patrick, Helen Margetts, Simon Bastow and Jane Tinkler (2006), *Digital Era Governance: IT Corporations, the State and E-Government*, Oxford: Oxford University Press.
- Dunleavy, Patrick, Christopher Gilson, Simon Bastow and Jane Tinkler (2009), 'The National Audit Office, the Public Accounts Committee and the risk landscape in UK public policy', The Risk and Regulation Advisory Council, available at: www.berr.gov.uk/deliverypartners/list/rrac/index.html.
- English, L. (2007), 'Performance audit of Australian public private partnerships: Legitimising government policies or providing independent oversight?', *Financial Accountability and Management*, **23** (3), 313–36.
- Fraser, S. (2007), 'Public scrutiny puts SAI audit practices to the test', *International Journal of Government Auditing*, January, 1–5.
- INTOSAI (2004), *Implementation Guidelines for Performance Auditing*, International Organisation of Supreme Audit Institutions, Vienna: INTOSAI.
- Ling, Tom (2007), 'New wine in old bottles? When audit, accountability and evaluation meet', in Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (eds), *Making Accountability Work*, New Brunswick, NJ: Transaction Publishers.
- Lonsdale, Jeremy (2000), 'Advancing beyond regularity: Developments in VFM methods at the National Audit Office 1984–1999', Unpublished PhD thesis, Brunel University.
- Lonsdale, J. (2007), 'Helping government to be responsive, responsible and respected: A reaction to Jocelyne Bourgon', *International Review of Administrative Sciences*, March, **73**, 31–6.
- Lonsdale, J. (2008), 'Balancing independence and responsiveness', *Evaluation*, **14** (2), 227–48.
- Lonsdale, J. and J. Mayne (2005), 'Neat and tidy . . . and 100% correct: Assuring the quality of supreme audit institution performance audit work', in John Mayne and Rob Schwartz (eds), *Quality Matters: Seeking Confidence in Evaluating, Auditing and Performance Reporting*, New Jersey: Transaction.
- Mulgan, G. (2006), *Good and Bad Power: The Ideas and Betrayals of Government*, London: Allen Lane.
- National Audit Office (2008), *Value for Money Handbook*, London: National Audit Office.
- National Audit Office (2010), *Progress with VFM Savings and Lessons for Cost Reduction Programmes*, London: National Audit Office.
- Nichol, Paul (2007), *Audit in a Democracy: The Australian Model of Public Sector Audit and its Application to Emerging Markets*, Aldershot: Ashgate.
- Office of the Auditor General (2008), 'What to expect – An auditee's guide to the performance audit process' at: www.oag-bvg.gc.ca/internet/English/meth_lp_e_30860.html (accessed 24 April 2010).
- O'Neill, Onora (2002), Reith Lectures 2002, Lecture 1 at: www.bbc.co.uk/radio4/reith2002/lecture1.shtml.
- Pollitt, C. (2006), 'Performance information for democracy: the missing link?' *Evaluation*, **12** (1), 38–55.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilka Summa and Marit Waerness (1999), *Performance Audit and Public Management in Five Countries*, Oxford and New York: Oxford University Press.

- Riksrevisionen (2009), *Annual Report of the Auditors General 2009*, Stockholm: Riksrevisionen.
- Riksrevisionen (2010), 'The role of the Swedish NAO' at: www.riksrevisionen.se/templib/pages/NormalPage_1768.aspx (accessed 25 April 2010).
- Roberts, Simon and Christopher Pollitt (1994), 'Audit or evaluation? A National Audit Office VFM study', *Public Administration*, **72** (4), 527–49.
- Schwartz, R. (1999), 'Coping with the effectiveness dilemma: Strategies adopted by state auditors', *International Review of Administrative Sciences*, **65**, 511–26.
- Seldon, Anthony (2009), *Trust: How We Lost It and How to Get it Back*, London: Biteback.
- Wilkins, Peter and Jeremy Lonsdale (2007), 'Public sector auditing for accountability: New directions, new tricks?', in Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (eds), *Making Accountability Work. Dilemmas for Evaluation and for Audit*, New Brunswick: Transaction Publishers.

Index

- Abma, TA 296, 297
academic studies 215–18
accountability 29, 36, 37, 42, 190,
209–10, 268, 298, 323–4
and efficient government 295
in evaluations 270–71
and learning 39–40, 268, 327–8,
332–3
National Audit Office 285
and performance 210–11
mechanisms 222–5
and performance improvement
studies 142
Supreme Audit Institutions (SAIs)
181
accountability approach 89–91
accountability mechanisms, public
sector performance 330
accountability mission, Supreme Audit
Institutions (SAIs) 199
accountability paradox 210–11
'Accountability and the Promise of
Performance' 210
accountability studies 142
added value 60
advocacy coalitions 187, 194, 202,
322
Africa 6
AFROSAI 6, 44
agency problems 83
Ahlbäck, S 23
Algemene Rekenkamer *see*
Netherlands, Court of Audit
Alon, G 177, 191, 199, 265
'ambush hearings' 240
American Evaluation Association 95
analysis 138–9
'analysis paralysis' 233
Anand, P 7
Anechiarico, F 217
anticipatory action 198
approaches, to performance auditing
314
argument
by induction 128
from an interviewee claim 128
from auditee claim 127
from expert opinion 127, 128
technique of VFM argument
construction 126–32
argument game 132
definition 125–6
argumentation schemes 126–31
arguments
deductive 124
'defeasible argumentation system'
120
episteme in 123–6
founded on evidence 118
framework for discussing 119–22
Argyris, C 224
Aristotle 119
on types of knowledge 122–3
ASAE 3000 (Australian Standard on
Assurance Engagements) 151
ASAE 3500 (Australian Standard on
Assurance Engagements) 150,
151, 156
assertions matrix 130–31
assessments 36, 43, 76
assurance, and standards 152–3
Astbury, B 243
Astill, S 9, 315
attention
for audited organisations 255
from other institutions 255–6, 263
attribution problem 237
audit
definition of 291
forms of 4
'gender audit' 232
increase in use 231–2

- use of term 135
 - versus evaluation 144
- audit criteria *see* norms
- 'audit explosion' 242, 313
- audit mandate 55–6
- audit results *see* results
- The Audit Society* 17
- 'audit society' 233, 238, 248
- audit topics *see* selection process
- audited departments, workload of 238
- auditee
 - financial cost for 200
 - perception and reactions 190, 193
 - regarding auditors' influence attempts 251
 - power relations with auditors 259
 - relationship with auditors 194
 - views on value of SAI's work 196, 197
- auditing
 - as ideal type 28
 - inclusion of performance auditing 29–34
 - meaning of term 26
- auditors 219–20
 - credibility 258–9
 - judgement 58
 - and performance improvement 211–15
 - relationship with auditee 194, 259
- Australia 7, 8, 14
 - ASAE 3000 151
 - ASAE 3500 150, 151, 156
 - assurance standards 153
 - Auditing and Assurance Standards Board (AUASB) 151
 - standard 156
 - Auditor General Act 151
 - Council of Auditors-General, Self Assessment and External Review Framework 162
 - development of performance audits 10
 - external review 162–3
 - National Audit Office (ANAO) 10, 159, 162, 198
 - Auditing Standards 150–51, 155–6
 - New South Wales Audit Office 158
 - Office of the Auditor General, Western Australia 162
 - public-private partnerships 241, 319
 - study on impact of performance audit 178
- Azuma, N 55, 154, 203, 204
- Baber, WR 216
- Balk, W 233
- banking crisis 2
- Bardach, E 284
- Barzelay, M 22, 25, 90, 91, 142, 143
- Bayesian theory 131
- Bechberger, E 108, 109, 120, 268, 282, 283, 284
- behavioural mechanisms 239–40
- Behn, R 15, 29, 39, 209, 210, 227, 295
- Belgium
 - Court of Audit (Rekenhof) 10, 53, 59, 175
 - Audit Manual 61
 - risk analysis indicators 60–61
 - selection process 62
- Bemelmans-Videc, M-L 16, 25, 39, 40, 52, 77, 148, 181, 231, 268, 270, 296, 297, 323
- Bennett, A 140, 240
- Bex, F 119, 125
- Black, R 268
- Blume, L 177, 190, 243
- Bohm, D 223
- Bolman, L 87
- Bongers, F 186, 189
- Booth 292
- Bouckaert, G 2, 109, 211, 218, 233, 253
- Bourn, J 268
- Bovens, M 66, 87, 271, 294
- Bowerman, M 7, 248, 272
- Boyd, K 282
- Bozeman, B 91
- Brans, M 186
- Brante, T 28
- Brauns, D 270
- Brazier, A 273
- Brown, R 177, 190, 191, 265
- Buber, M 223
- Bucuvalas, M 186, 249, 253
- Bunting, M 238
- Burke, B 313, 315, 316
- Caiden, GE 220
- Callahan, K 325

- Campbell, S 184, 186
- Canada 7, 189–90
- Fraser Institute 233
- legislation 8–9
- Office of the Auditor General (OAG) 7, 159, 198, 212, 317
- public dialogue 222
- Results for Canadians 3
- studies on impact of performance audit 179–80
- case studies 101, 195
- chance events 195, 203
- change 188, 237
- in environment 318
- characteristics, performance auditing 6–9
- Chelimsky, E. 24, 35, 66, 89, 136, 142, 143
- Cinderella subjects 60
- circular causality 263–4
- Clemens, J 234
- cognitive biases 240
- cognitive coupling 253, 261
- cognitive-behavioural mechanisms 242
- coherence, judgements 138
- Commonwealth of Australia 162
- communication 185, 193, 318
- complexity 293, 295, 296, 318
- compliance auditing 219
- conclusions
- challenging 125–6
- evidential support for 121–2, 123–4
- construction, of performance auditing studies 143
- constructive argumentation 296–7
- consultation 241
- consultative methods 102–103
- contribution, of performance auditing 320–26
- Cooper, D 145, 271
- correspondence judgements 138, 143
- corruption 190, 233
- Court, J 186
- credibility
- of auditors 258–9
- government 330
- Creerar, L 243, 313
- Crijns, M 237, 239
- criteria 294
- criticism, of performance auditing 209–10, 218–21
- Cronbach, L 140
- cross-government perspective 328, 332
- cross-government studies, choice of methods 112
- cultural influences 156–7
- cutbacks, discretionary activity 2
- Dahanayake, SJ 241
- Dahl, R 325
- data, on impact 321
- data analysis 102, 116
- data collection, innovative techniques 242
- ‘data dumping’ 294
- De Lancer Julnes, P 3, 324
- De Vries, GJ 82, 177, 189, 192, 193, 194, 250, 253, 265
- Deal, T 87
- debate 257, 332, 333
- decision analytical modelling 102
- decision making process 183
- ‘decoupling’ 200
- deductive arguments 124
- deductive and defeasible logics 123–5
- ‘defeasible argumentation system’ 120, 122, 132
- attacks on defeasible argument 125
- defeasible logic 133
- research 125
- deficit reduction plans 2
- definitions, of performance audit 1, 6–9, 34–7, 43, 218–19, 257, 291
- delivery, of policy measures 301
- democracy, and performance auditing 322–6
- democratic citizenship 332
- Dempster-Shafer theory of evidence 131
- development, of performance auditing 4, 312
- diffusion 185
- ‘digital era governance’ 316
- discretionary activity, cutbacks 2
- Dixon, NM 222
- document examination 100, 102, 115
- Dolmans, LJJ 204, 238, 248, 253, 265
- Donaldson, S 41
- Dryzek, JS 296

- Dubnick, M 209, 210, 211, 222, 227, 315
- Dunleavy, P 313, 316, 319
- 'dynamic norms' 291
- economic crises 319
- Edelenbos, J 51, 52, 271
- effectiveness 51–2, 91
- norms ('audit criteria') 77
- and recommendations 283
- effects, of performance audits 182–3, 188
- efficiency savings programmes 2
- Eichenberger, R 203, 242, 265
- Elliott, J 204, 238, 248
- Elster 243
- embezzlement 30
- English, L 241, 319
- environment, changes in 318
- environmental factors 260
- episteme, of rigour in performance audit argument 123–6
- Estonia, study on impact of performance audit 177–8
- Etverk, J 177, 189, 190, 193, 253
- European Court of Auditors 196
- evaluation 23, 25, 35–6, 41–2, 51, 66, 101–102, 295, 316
- in accountability settings 270–71
- choice of type 85
- impact of 181, 184
- influence of 181
- learning from 274
- outside accountability 271
- and performance auditing 136
- priorities 184
- programme 142
- realistic 142
- responsive 297
- versus audit 144
- Evaluation and Auditing* 23
- evaluation culture 194
- evaluative activity, and learning 269–70
- Evans 292
- evidence 95, 140
- about performance 316–17
- argument founded on 118
- Dempster-Shafer theory of evidence 131
- interviews as 130
- iterative approach to structuring 13
- methods of collection 100–105
- quantifying evidential support 131
- and standards 152–3
- tendency to both support and refute a proposition 127
- use of 118
- weighting evidence and audit methods 131–2
- evidence based performance auditing 241–2
- expert panels 102
- external quality review
- criteria 164
- ex-ante, formative 167–8
- ex-post, summative 167
- overall assessment 164, 165
- of reports 166
- case study 164–5
- scoring range 164
- external review 158, 162–3, 317
- fairness 147
- feedback 298, 306
- Fenger, HJM 77
- financial audit, definition 121
- financial impact of audits 196, 198
- Finer, H 30, 31
- Finland 4
- fiscal policy 177
- Fisher, F 296
- fitness for purpose 158
- Flint, D 28, 29, 95
- focus groups 101, 102
- Forester, J 296
- FrameWorks Institute 67
- France, Cour des Comptes 9
- Fraser, S 219, 222, 313, 316
- Frederickson, H 210, 223
- Freeman, H 274
- Freeman, R 284
- Frey, BS 204
- Friedberg, A 215
- Friedrich, CJ 30
- Frissen, PHA 293
- Fundamental Auditing Principles and Auditing Guidelines 148
- Funkhouser, M 192, 209, 213, 214, 216, 218, 223, 224, 236

- Funnel, W 91
 Furubo, JE 25, 118, 136, 285, 286, 311
 'gaming' 239
 'gender audit' 232
 Gendron, Y 145, 241, 271
 Geographical Information System (GIS) 102
 George, AL 140, 240
 Germany 38
 Glynn, JJ 89
 goal-free evaluation model 297
 good practice 268, 277, 332
 and performance auditors 282–4
 procedural advice 282–3
 government
 effectiveness 177, 236, 302
 examination by citizens 325–6
 innovation and renewal in 238
 learning in 279–81
 quality of 320
 trust in 221, 330
 'government failures' 233
 government performance
 impact of performance audit on 320–22
 and performance audit 190
 Gray, A 25, 231
 Gronlund, A 7
 Guba, EG 296
 Gunvaldsen, J 9
 Gupta, V 236
 Guthie 23

 Halachmi, A 211
 Halligan, J 2, 218, 314
 Hammond, K 137, 139
 Hanney, S 186
 Hansard 114
 Hansen, HF 242
 Harper, R 139
 Harrison, M I 87
 Hatherly, D J 178, 194, 195, 198, 265
 Havens, H 71
 Hawthorne-effect 189
 Haynes, W 313, 315, 316
 Haynes Watts 198
 health studies, choice of methods 111
 Hedstrom, P 243
 Henry, GT 240

 Hodgkinson, GP 91
 Hollingsworth, K 15
 Holzer, M 325
 Hood, C 91, 231
 House, E 66
 Howe, K 66
 Howlett, M 269
 Hughes, EFX 249, 253

 'ideological capture' 241
 impact 176, 176–87, 233–7
 at local government level
 multiple case study 248–67
 average scores 256
 difficulties and limitations 262–4
 interview selection 254–5
 research approach 250–55
 research findings 255–62
 results by category of
 performance indicator 257
 results by category of success
 factors 260
 selection criteria 254
 on audited organisation 251
 causality 263
 and chance events 195
 conceptual 180, 182, 198
 impact studies on 188
 data on 321
 definition in impact studies 187,
 189–90, 265
 design of reviewed studies 177–80
 empirical research on 187–201
 facilitating factors 184–7, 192–5
 in impact studies 193
 macro level 184, 186, 193
 meso level 184, 186, 193
 micro level 184, 185, 193
 factors determining 202
 forms of 180–81
 on government performance 320–22
 indicators used by SAIs 197–8
 indirect, studies 248
 information in reviewed studies 191
 instrumental 180, 182, 189, 190, 199
 impact studies on 188
 interactive 180, 183
 impact study on 188, 189
 maximising 322
 meaning of 176, 237

- measurement 249, 331
- non-financial 235
- political-legitimizing 180, 183
 - impact studies on 188
- qualitative 235
- quantifiable financial 235
- quantitative information 190–92
- tactical 181, 183, 189
- types 181, 182–3
 - in impact studies 188
 - see also* side effects
- impact assessments 303–304
- independence 33–4, 37, 51, 147, 152, 241, 273
 - of information in audit reports 326
- indicators, advantages and risks of 296
- indirect impacts *see* side effects
- inductive reasoning 124–5
- influence attempts
 - in performance auditing 251
 - success factors in 252, 258
- information, in audit reports, independence 326
- Ingraham, PW 178, 190
- innovation 113, 331
- Innvaer, S 186
- institutional setting, performance auditing 9
- In't Veld, RJ 294
- integrated theory of influence 249
- intellectual depth, of audits 239–40
- internal auditing 224
- Internal Auditor* 224
- internal quality assurance procedures, and peer review 162
- internal validity 140
- International Auditing and Assurance Standards Board (IAASB) 151, 154
- international comparisons 101
- International Federation of Accountants 151
- International Monetary Fund 139
- International Organization of Supreme Audit Institutions *see* INTOSAI
- international performance audit standards 148
- International Standard on Assurance Engagements 3000 (ISAE 3000) 151, 154, 157
- International Standards on Auditing 132, 154
- International Standards Organisation, quality management systems 158
- interviews 100, 102, 110
 - research 144
 - use as evidence 130
- INTOSAI 26, 58, 67, 148, 154, 157, 291, 292, 298, 316, 318
- Auditing Standards 6–7, 75, 149–50, 153, 161
 - guidelines 26–7, 38
 - The Implementation Guidelines for Performance Auditing* 213, 219, 291
- Performance Audit Subcommittee 154
- performance measurement within 196
- intuition 138–9
- Ireland
 - external quality review of reports, case study 164–5
 - Office of the Comptroller and Auditor General 160, 164, 317
 - peer review 159
 - case study 160–61
- Israel, studies on impact of performance audit 177, 179, 191
- ISSAI 3000 148
- Jackson, A 238, 265
- Jacobs, JB 217
- Jenkins, B 231
- Johnsen, A 203, 233, 265
- Johnson, K 186
- Johnston, WP 178, 191, 194, 195, 198, 265
- judgements
 - in audit work 135–8
 - coherence 138
 - correspondence 138
 - panoptic 138
 - in performance audit 136
 - validity of statements 137–8
- Julnes, G 240
- Karlsen, R. 9
- Keen, J 23, 108, 118

- Kemp, R 226
 King, E 284
 Kirkhart, KE 181, 249
 Klay, WE 218
 knowledge, Aristotle on types of
 122–3
- Landers, B 240, 272
 Lane, JE 39
 Lapsley, I 4, 135, 178, 191, 234, 265
 learning 182, 188, 291, 306, 320
 and accountability 39–40, 268,
 327–8, 332–3
 in an accountability setting 270–71
 and evaluative activity 269–70
 from evaluation 274
 in government 279–81
 government 269
 improving role of performance audit
 331–2
 and National Audit Office 271–81
 organizational 223–4
- Leeuw, B 242, 265
 Leeuw, FL 15, 25, 200, 204, 224, 232,
 233, 236, 237, 238, 239, 241, 242,
 243, 248, 253
- legislation
 Canada 8–9
 flexible 328
 in selected countries 5
- Lehtonen, M 142, 268
 Leviton, LC 186, 249, 253
 Levitt, M 137
 Levitt, R 137
 Lima Declaration 67
 Lincoln, YS 296
 Lindquist, EA 186
 Ling, T 25, 105, 113, 137, 297, 319
 Lipsey, D 272
 Lipsey, M 274
 literature, on performance auditing
 311
 literature reviews 102
 local government, impact of
 performance audit, multiple case
 study 248–67
Local Government Auditing Quarterly
 224
 logics, deductive and defeasible
 123–5
- Lonsdale, J 4, 9, 16, 24, 25, 39, 40, 51,
 52, 54, 97, 98, 99, 102, 105, 114,
 118, 135, 147, 157, 158, 163, 181,
 219, 231, 235, 240, 268, 270, 272,
 279, 297, 313, 315, 317, 323
- Lutrin, C 215
- McBride, T 109
 macro-economic indicators 190
 Madan, P 91
 Mahoney, J 240
 Majone, G 85, 86, 89, 296, 297
 malfeasance 31
 managerialist-consumerist versus
 political approach 83–5
- MAPE (Means Activities Performance
 Effects) spectrum 196
- Mark, MM 33, 240
 Marlowe, J 226
 Martin, S 231, 242, 271
 Massey, A 270
 Mayne, J 16, 23, 25, 97, 147, 157, 158,
 163, 233, 279
- Mazur, J 67
 MEANS grid 317
 means-end reasoning 127
 media interest 183, 188, 193, 194, 196,
 197, 256, 264, 313
- Meklin, P 203
 methodological quality, of audits 240
 methodology 315
 in peer review 160
- Methodology Square 138–43
 evaluative activity 142–3
 and performance auditing 140–42
- methods
 choice of 98–9
 constraints 108
 in cross-government studies 112
 environmental factors 113–15
 existing knowledge 108
 health studies 111
 nature of subject to be examined
 105–106, 109
 organisational factors 110–13
 previous use of method 108
 project-centred factors 109–10
 team-centred influences 105,
 108–109
 factors influencing choices 105–15

- number used per study 104
- trends in use 1908s–2000s 99–105
 - 1986–88 100–101
 - 1992–95 101
 - 1997–98 101
 - 2003–05 101–102
 - 2005–09 102, 105
- weighting evidence and audit methods 131–2
- Meyer, M 236
- Michael, B 210
- ‘micro issues’ 137
- Milgrom, T 179, 191, 195, 199, 249, 265
- Mintzberg, H 86
- ‘mission creep’ 332
- modelling 102
 - National Audit Office 108–109
- monitoring process, Netherlands 64
- monitoring systems 54
- monotonic deductive reasoning 124
- Morgan 87
- Morgan, S 67, 68
- Morin, D 23, 90, 91, 179, 189, 190, 191, 192, 194, 195, 199, 200, 201, 202, 248, 249, 250, 253, 255, 258, 262, 264
- Mossberger, K 284
- Moynihan, DP 178, 190
- Mulgan, G 14, 325
- Mulgan, R 91
- Munro, E 204
- Myrdal, G 41

- Nalbandian, J 225
- National Systems for Accountability, Transparency, and Oversight of Economic Activities* 216
- Neely, A 238, 265
- Netherlands 6, 8, 10, 12, 232
 - Accountability Day 300, 306
 - analysis of societal problems 64
 - children, participation in sport 302
 - Court of Audit 6, 8, 10, 12, 15, 38, 62–3, 75, 159, 161, 162, 175
 - audit on Ecological Structure in the Netherlands 300
 - audit on energy savings policy programme 299–300
 - audit on jobless people 304
 - focus on policy level 84
 - ‘good management’ norm 79–80
 - ‘good policy’ norm 78–9
 - ‘good results’ norm 79, 80
 - Information Communication Technology 82
 - norms 77–81
 - official sources of norms 81
 - reports on credit crisis 319
 - responsiveness in performance audits 298–303
 - WWWHW (Why Who What How When) criteria 162
 - effective influencing 61
 - female professors at universities 302
 - integrated risk analysis 64–5
 - issue matrix 64
 - Ministry of Finance, policy audits 303
 - monitoring process 64
 - peer review 159, 161
 - policy areas 63–4
 - recurring audits 63
 - Scientific Council for Government Policy 292, 295
 - selection process 62–6
 - stimulating rail transport 301
 - strategic choices 63
 - strategy, planning and selection tools 65
 - studies on impact of performance audit 177, 180
 - see also* impact, at local government level, multiple case study
- New Public Management 241, 257–8, 262, 263, 282
- New Zealand
 - Annual Plan of the Controller and Auditor-General 221
 - peer review 159
- Nichol, P 7, 8, 313
- non-compliance 156
- ‘non-monotonic’ logics 119
- Noordegraaf, M 86
- norms 75–6
 - Algemene Rekenkamer 77–81
 - development of 76–7
 - effectiveness 77
 - ‘good management’ practices 77, 79
 - ‘good policy’ 77, 78–9

- ‘good results’ 79, 80
- National Audit Office 81
- official sources 88
- official versus pluralist approach 88–9
- ‘output/service delivery’ 77
- performance accountability of the government towards the parliament 77–8
- reports addressing certain types of 78
- Norway 9
- Nutley, S 137, 176, 186
- objectives 294
 - advantages and risks of 296
- objectivity 147
- O’Connell, AJ 54
- OECD 25, 268
- Oh, CH 186
- Ollerearnshaw, S 284
- O’Neill, O 324
- ‘open target spending’ 302
- organizational learning 223–4
 - definition 224
- O’Shaughnessy, K 236
- ‘ossification’ 233, 238, 265
- outcome/impact, versus reliability 40–42
- outcomes 319
 - anticipation of 182, 188, 189
- Overman, S 282
- Page, E 9, 268, 282, 283, 284, 315
- panoptic judgements 138
- Parker 23
- Parker, LD 178, 195, 198, 265
- parliament 194, 196
 - discussion of SAI’s audit reports 197
 - view on value of SAI’s work 197
- ‘partnerial’ investigations 241
- Patton, MQ 34, 51, 142, 181
- Pawson, R 140, 141, 240
- peer review 158, 159–62
 - case study 160–61
 - engagement with the value-for-money client 160
 - and internal quality assurance procedures 162
 - methodology 160
- New Zealand 159
 - scope 161
 - staff skills 161
 - and standards 161
 - topic selection 160
- performance
 - and accountability 210–11
 - mechanisms 222–5
 - meaning of word 75
- performance auditing
 - academic studies 215–18
 - accountability, and learning 332–3
 - accountability role 327
 - adoption in the US 217
 - approaches to 314
 - characteristics 6–9
 - choice of most valid type 87
 - contribution of 320–26
 - criticism of 209–10
 - cumulative effect 263
 - defining 330–31
 - definitions 1, 6–9, 34–7, 43, 257, 291, 326
 - INTOSAI 219
 - lack of clarity in 218–19
 - and democracy 322–6
 - for ‘democracy’ or ‘improvement’ 89–91
 - development 4, 312
 - Australia 10
 - downsides of 331
 - and evaluation 136
 - evidence based 241–2
 - evolution 6
 - external review 162–3
 - forms of 7
 - future of 330–33
 - and government 175–6
 - growth in scale and intensity 312–13
 - as ideal type 23, 44
 - ideas and concepts 25–6
 - impact 176–87, 233–7
 - on government performance 320–22
 - measuring 331
 - importance of 232
 - inclusion in auditing 29–34
 - influence attempts in 251
 - institutional setting 9, 22
 - interpretations of 26–7

- learning, improving role in respect of 331–2
- literature 24–5, 311
 - responsiveness in 296–8
- managerialist–consumerist versus political approach 83–5
- and the Methodology Square 140–42
- nature and meaning of 326–30
- norms used in 75
- peer review 158, 159–62
- performance indicators for 251
 - as policy analysis and evaluation 15
 - and politics 38–9, 44
 - preventive effects 264
 - professional basis 157
 - and public sector performance 4–9
 - quality arrangements 97
 - quality assurance arrangements 312
- reports
 - publication 14
 - quality 157
- responsiveness in 289–307
 - improving 332
- role of 11, 14–16
- side-effects of 200–201
- as a social influence process 250
- statutory powers 312
- studies, construction of 143
- theories underlying 237
- theory and practice, improving 331
- and trust 324–5
- value of 201, 321
- performance auditors 9
 - and good practice 282–4
- performance audits
 - effects of 182–3
 - quality of 147
 - side-effects 237–9
 - standards 147
- performance improvement, and auditors 211–15
- performance indicators 158
 - auditees' perception 257
 - contribution to public debate 257
 - impacts generated by audit 257
 - for performance audit 251, 253
- performance management 294–6
 - and results-oriented budgeting 305
- performance measurement 2
- 'performance paradox' 233, 238
- performance problems 82
 - accountability 87–8
 - explanation 87
 - exposing versus explaining 85–8
- Perrin, B 16, 181, 268, 270, 297, 323
- Peters, BG 211, 236
- Philosophy and Principles of Auditing* 28
- phronesis 132–3
- Pierce, PA 243
- policy audits 303
- policy disasters or failures 328
- policy evaluation 176, 269
- policy evaluation literature, responsiveness in 296–8
- policy implementation 294
- political responsiveness 225
- political taboos 58
- politics 67
 - and performance auditing 38–9, 44
- Pollitt, C 1, 4, 15, 16, 22, 24, 38, 39, 41, 53, 75, 76, 77, 86, 97, 109, 116, 136, 142, 234, 238, 241, 249, 253, 264, 265, 272, 273, 282, 311, 315, 332
- Pollock, J 125
- Pong, CKM 178, 191, 234, 265
- Power, M 4, 17, 28, 135, 204, 231, 242, 295
- Prakken, H 119, 125
- Preskill, H 269
- preventive effects, performance auditing 264
- principle–agent relationships, public audit system as a chain of 83–4
- procedural advice 282–3
- process 185, 260
- process requirements, unwarranted 156
- professional basis, of performance auditing 157
- programme evaluation 142
- Prosser, M 232
- public audit system, as a chain of principle–agent relationships 83–4
- Public Audit Forum 240
- public debate, contribution to 251
- public dialogue 222–3
- public employment, reduction 2
- public expenditure, value from 330
- public funds 53
- public policy 289

- as an audit topic 292–3
 - complexity 318
- public resources 53
- public sector performance
 - accountability mechanisms 330
 - and performance auditing 4–9
- Put, V 23, 77, 236, 239, 248, 249
- Pyper, R 270
- qualitative data 113
- quality 97, 157–9, 315, 316
 - of audit process 157
 - of audit products 157
 - definitions and interpretations 157
 - of government 320
 - implications of standards 155–7
 - performance audit reports 157
 - of performance audits 147
 - of reporting/reports 147, 149–50, 158–9, 192
 - technical 261
- quality assurance 157–8, 312
- quality control 158
- quality review
 - options for 163, 165
 - ex-ante*, formative external review 167–8
 - ex-post*, summative external review 167
 - reliance on internal management and supervision (do nothing approach) 165–7
- quantitative data analysis 102
- questionnaires 101
- Raaum, R 67, 68
- Radcliffe, VS 77, 189
- Ram, V 273
- Ramesh, M 269
- RAND 17
- Rawnsley, A 1
- realistic evaluation 142
- reasoning, inductive 124–5
- recommendations 185
 - acceptance 182, 188
 - approach of standards to 153–4
 - and effectiveness 283
 - implementation 177, 182, 188, 196, 197, 198, 263
 - timeframes for implementation 199
- Reed, C 128
- reform projects 194
- regulation 2, 8–9
- ‘regulatory capture’ 330
- Rekenhof *see* Belgium, Court of Audit
- relevance 261
- reliability, versus outcome/impact 40–42
- reports 185
 - external quality review, case study 164–5
 - independence of information in 326
 - quality of 147, 149–50, 158–9, 192
 - self-initiated review 163
 - structure, testing of inferences 129
- research 53–5, 71, 82, 140
 - interviews 144
 - method-driven and theory-driven 140
- resistance 239
- responsive evaluation 297
- responsiveness
 - in performance auditing 289–307
 - improving 332
 - Netherlands Court of Audit 298–303
 - audit process 298, 298–300
 - choice of actions 299–300
 - follow up reports 298
 - impact assessments 303–304
 - ‘reality checks’ 300–303
 - responsive audit approach 299
 - selection of themes and topics 298
 - transparency 298–9
 - trust 299
 - in political evaluation and performance audit literature 296–8
- results 181, 182, 185, 187
 - dissemination 203
 - timeliness 194
- Rethinking Democratic Accountability* 29, 209
- Rieper, O 242
- rigidities 200
- Riksrevisionen *see* Sweden, National Audit Office (NAO)
- risk analysis 57, 60
 - assumptions about 69–70

- design and implementation of
 - systems 68–9
 - evidence base 68–70
- risk matrix 66
- Rist, R 25
- RMO 233
- Roberts, NC 222
- Roberts, S 315
- Rogers 216
- role of performance auditing 11, 14–16
- Ross, WD 118
- Rossi, P 270, 274
- Rossi, PF 186
- Rossi, PH 237

- Saint-Martin, D 90, 91, 202
- Sartori, G 26
- 'savings for the citizens' 235
- Schaltegger, CA 203, 235, 265
- Scharaschkin, A 120, 135, 137
- Schelker, M 180, 190, 194, 203, 243, 265
- Schum, D 131
- Schwandt, TA 294, 296
- Schwartz, R 24, 32, 41, 77, 78, 147, 179, 191, 195, 199, 221, 233, 249, 265, 315
- Scotland 13–14, 314
 - Audit Scotland 317
- Scott 87
- Scriven, M 35, 297
- scrutiny 231
- scrutiny committees 15
- 'secondary dissemination' 274, 276
- Seldon, A 324
- selection process 53–5
 - Belgium 62
 - criteria 54
 - 'financial importance' 67
 - institutional level 55–6
 - models of 54–5
 - Netherlands 62–6
 - organizational level 56–8
 - risk analysis 57, 60
 - selection as a value-laden process 66–8
- self-initiated review, reports 163
- Serna, A 204
- Shafer, G 131
- Shand, D 7

- Sharma, N 23, 81, 96, 114, 137, 273
- Shirom, A 87
- Shore, C 248
- short termism 238, 265
- shortcomings, in audits 199
- 'shower-of-shot-approach' 301
- side-effects
 - 'decoupling' 200
 - narrowing of vision 200
 - occurrence of 239–41
 - ossification 238
 - of performance audits 237–9, 261, 262
 - literature 200–201
 - unintended and undesired 237–9
 - reasons for 201
 - resistance 239
 - rigidities 200
 - short termism 238
 - suboptimisation 200
 - tunnel vision 238, 265
- Siemiatycki, M 82
- Sigma study 53
- site visits 101, 102
- skills 113, 115, 161, 297
- Smith, P 238, 265
- social mechanisms approach 210
- Solesbury, W 137
- Sonnischen, R 224
- Spirit of Public Administration* 223
- The Spread of Performance Auditing Among American Cities* 216
- Stake, R 297
- stakeholder consultations 102, 110, 183, 188
- stakeholders
 - advocacy coalitions 187, 194
 - influence of 314
 - information from 306
- Stame, N 25
- standards 132, 316
 - approach to assurance 152–3
 - approach to evidence 152–3
 - approach to recommendations 153–4
- Australia 150–51
- codifying 317
- comparison 149
- and consistency 148
- implications for quality 155–7

- merits of different approaches 154–7
- and peer review 161
- of performance audits 147, 148–52
- UK, National Audit Office 150
- US, Government Accountability Office (GAO) 150
- Starkey, K 91
- state structure 55–6
- ‘static ambitions’ 294–6
- statistics 35
- status quo, challenges to 261, 262
- Sterck, M 54, 236
- stock markets, capitalization 1
- Stone, D 284
- ‘strategic audit failures’ 84–5
- strategic choices 66–7
- Strathern, M 135
- Straussman, J 91
- Streim, H 83
- studies
 - accountability and performance improvement studies 142
 - types and purposes 144
- Studies in State Audit* 215
- study design
 - and choice of methods 98–9
 - trends in use of methods 1980s–2000s 99–105
- study selection process 51–2
- style of audits 240
- suboptimisation 200
- Summa, H 15, 98, 136, 142
- Supreme Audit Institutions (SAIs) 4, 7–8, 22–3, 52, 147, 311
 - accountability 181
 - accountability mission 199
 - characteristics 190
 - core activities 330
 - criteria and norms 239
 - differences in instructional and organizational structures 194
 - impact of 201–202
 - importance of quality of information in reports 158–9
 - limits and constraints 328–30
 - models 9
 - organizational structure 58–9
 - performance indicators 196
 - performance measurement data 196–200
 - priorities and decisions 53–4
 - and quality 316
 - quality assurance systems for financial audits 157
 - reduced budgets 314
 - task of 175
 - types of impact indicators used by 197–8
 - use of norms 76–7
 - variation of standards 151–2
- surveys 101, 102, 109
- Sutcliffe, S 186
- Sutherland, SL 241
- ‘swamping’ 239
- Swedberg, R 243
- Sweden 4, 14, 326
 - National Audit Office (NAO) 7–8, 12
 - ‘State Guarantees in the Financial Crisis’ 319
- Switzerland 235
- Talbot, C 2, 203
- ‘taming trap’ 305
- target groups 301
- targets 294
- Taut, S 270
- technique, of VFM argument construction 126–32
- technical quality 261
- t’Hart, P 87
- theories, underlying performance auditing 237
- theory and practice, performance auditing 331
- thinking, intuition and analysis 138–9
- third parties, role of 194
- Tilley, N 140, 141, 240
- timing, of performance auditing 259
- toolkits 277, 282, 315
- topic selection 53–5, 305
- Torgler, B 203, 235, 265
- Torres, R 269
- Toulmin, S 133
- Townley, B 145, 271
- Tranfield, D 91
- transparency 71, 221, 234, 298, 324, 333
- trust 43, 227, 299, 324
 - in government 221, 226, 330

- and performance auditing 324–5
 - on the work floor 238
 - tunnel vision 238, 265
 - Turksema, R 12
 - typology of knowledge, for value for money audit 122–3
- UK 4, 9
- Audit Commission 13
 - Audit Scotland 14
 - Committee of Public Accounts 15, 272, 280
 - Comprehensive Spending Review 3
 - District Audit 13
 - Equity Commission 232
 - Gershon Review 3
 - National Audit Act 97–8
 - National Audit Office 7, 9, 12–13, 14, 97, 167, 175, 318
 - accountability role 285
 - accountability structures 96
 - additional products and summaries 276
 - audit on privatization 235
 - audit reports, discussion of findings 277
 - ‘clearance’ process 81, 125, 137
 - ‘cold reviews’ 163, 167
 - and Committee of Public Accounts 280
 - Consultancy Assessment Tool 277
 - effects of 232
 - focus on management level 84
 - hospital acquired infection report 281
 - ‘hot reviews’ 167–8
 - and learning 271–81
 - evidence for 272–3
 - material on good practice 277
 - methods used in reports 100
 - modelling 108–109
 - norms 81
 - parliamentary context 114
 - principles and quality thresholds 151
 - recommendations 281
 - references to learning in annual reports 275–6
 - report
 - on credit crisis 319
 - ‘End of Life Care’ 109
 - Foot and Mouth disease 281
 - on government saving programmes 325
 - ‘Making a Difference. Performance of maintained secondary schools in England’ 81
 - ‘Managing Attendance in the Department of Work and Pensions’ 121
 - on nationalisation of the banks 319
 - ‘Reducing Vehicle Crime’ 127, 128
 - ‘Stroke Services’ 141–2, 143
 - ‘Supporting people with autism through childhood’ 137
 - ‘Tackling Pensioner Poverty’ 82
 - self-initiated review of reports 161
 - sources of norms 81
 - standards 150
 - and evidence 152
 - for recommendations 153
 - Strategy 284, 285–6
 - study design 98
 - target ratio of benefit to cost 214
 - toolkits 277
 - value for money audit 96–8
 - Value for Money handbook 150
 - value for money standards 317
 - Operational Efficiency programme 3
 - Public Accounts Committee 81, 96
 - public spending cuts 2
 - Public Value Programme 3
 - study on impact of performance auditing 178
 - Treasury 3, 232
 - Value for Money Update 3
 - Wales Audit Office 14
- US
- American Recovery and Reinvestment Act 1
 - anti-corruption efforts 217
 - City Auditors Office, Kansas City, Missouri 221
 - City of San Jose, Office of City Auditor 214
 - Environmental Protection Agency (EPA) 212, 214, 219

- Government Accountability Office (GAO) 13, 54, 55, 212, 214, 240, 312, 316
- GAGAs (generally accepted government auditing standards) 150, 316
- use of their party data 325
- Government Auditing Standards 132, 212, 235
- Government Performance and Results Act 3
- infrastructure 226
- legislative oversight offices 234–5
- National Association of Local Government Auditors 214–15
- performance audit adoption by year 217
- Portland Office of the City Auditor 213
- Post Audit, Kansas Legislative Division 212–13, 220
- studies on impact of performance audit 177, 178, 180, 191
- Vaessen, J 237
- validity, internal and external 140
- value
- added 60
 - of performance auditing 201, 321
- value for money 3
- technique of argument construction 126–32
- value for money audit
- methods 115
 - National Audit Office 96–8, 110
 - standards 152
 - summative judgements 136
 - outcomes 12
 - typology of knowledge for 122–3
- value for money studies, building 143–4
- Van der Knaap, P 12, 269, 295, 296, 297, 298, 306
- Van der Meer, FB 51, 52, 180, 189, 192, 250, 265, 271
- Van Der Straeten, T 186
- Van Loocke, E 249
- Van Thiel, S 204, 233, 236, 243
- Vanlandingham, GR 234, 235
- Varone, F 184, 186
- Vedung, E 35, 88
- Verheij, B 125, 126
- Voigt, S 177, 190
- Vreeswijk, GAW 125
- Wales 13–14
- Walker, DM 212, 226
- Walsh, AH 240
- Walsh, K 282
- Walton, D 128
- Webb, J 176
- Weber, M 28
- Weets, K 203, 204, 234, 243, 265
- weighting evidence and audit methods 131–2
- Weiss, C 51, 66, 86, 176, 186, 249, 253, 271
- ‘What Public Bodies can Expect from their Auditors’ 240
- ‘What to Expect: An Auditee’s Guide to the Performance Audit Process’ 317
- Wheat, E 216, 217
- White, F 15
- Wholey, JS 86
- ‘wicked problems’ 220
- Wiggan, J 203
- Wildavsky, A 223
- Wilkins, P 51, 231, 268, 272, 313
- Wilkinson, J 231
- Willemsen, F 242
- Williams, R 232
- Willing Slaves* 238
- Wisler, C 23, 136
- Wolman, H 284
- Working Group on Value and Benefits of SAIs 204
- workload, of audited departments 238
- World Bank 176, 196
- Wright, S 248
- WRR (Netherlands, Scientific Council for Government Policy) 290, 293, 295
- Zacchea, N 216, 218

