INTENTION AND AGENCY

PHILOSOPHICAL STUDIES SERIES
IN PHILOSOPHY

VOLUME 33

DONALD F. GUSTAFSON

Department of Philosophy, University of Cincinnati

INTENTION AND AGENCY

D. REIDEL PUBLISHING COMPANY

A MEMBER OF THE KLUWER  ACADEMIC PUBLISHERS GROUP

DORDRECHT / BOSTON / LANCASTER / TOKYO

Library of Congress Cataloging-in-Publication Data

Gustafson, Donald F., 1934-
Intention and agency.

(Philosophical studies series in philosophy ; v. 33)
Bibliography: p.
Includes indexes.
1. Act (Philosophy) 2. Intentionalism. 3. Causation. I. Title. II. Series.
B105.A35G87 1985 128'.4 85-25623

ISBN-13: 978-94-010-8514-4 e-ISBN-13: 978-94-009-4520-3
DOI: 10.1007/978-94-009-4520-3

Published by D. Reidel Publishing Company,
P.O. Box 17, 3300 AA Dordrecht, Holland.

Sold and distributed in the U.S.A. and Canada
by Kluwer Boston Inc.,
190 Old Derby Street, Hingham, MA 02043, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers Group,
P.O. Box 322, 3300 AH Dordrecht, Holland.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	vi
CHAPTER ONE: Introduction	1
CHAPTER TWO: Passivity and Activity in Intentional Action	20
CHAPTER THREE: Intending, Judging, and the Cognitive Model	36
CHAPTER FOUR: Further and Future Intentions	53
CHAPTER FIVE: Expressions of Intention	69
CHAPTER SIX: Intentions: The Structures	84
CHAPTER SEVEN: Intentions: The Contents	105
CHAPTER EIGHT: Dynamics of Intentional Action	122
CHAPTER NINE: Agency and Psychological Causation	148
CHAPTER TEN: Questions and Some Answers	166
CHAPTER ELEVEN: Activity and Development	191
CHAPTER TWELVE: Overview	199
INDEX OF NAMES	206
INDEX OF SUBJECTS	207

ACKNOWLEDGEMENTS

Chapter Two of this work previously appeared in *Mind,* Volume XC, in January of 1981. The first sections of Chapter Five appeared in *Mind,* Volume LXXXIII in 1974, July. *Inquiry,* Volume 18, 1975 contains an essay of mine with the title 'The Range of Intentions'; some of the ideas in this essay are now in Sections 1–4 of Chapter Six. These three essays were first written with aid from the Charles Phelps Taft Fund. Additionally, the Research Council of the University of Cincinnati provided funds for final preparation of this manuscript. Irene Diehl and Mary Von Nida were very helpful at various stages of the manuscript production.

One of the admirable features of all of the writings of Gilbert Ryle is the absence of annoying footnotes. Ryle never used footnotes. I have followed his practice. Parenthetical remarks (as alternatives to footnotes) are annoyance enough.

My deepest thanks go to my wife, Penny Freppon, for her enthusiastic support. This work is dedicated to those students at several universities who listened to these ideas expressed and responded to them.

CHAPTER ONE

INTRODUCTION

> The powers of seeing, hearing, remembering, distinguishing, judging, reasoning, are speculative powers; the power of executing any work of art or labour is active power.
>
> Thomas Reid

I

Some causal efficacy is due to persons. And, some of the causal efficacy due to persons is imparted by, not merely to, them. Further, some of the causal efficacy due to persons and imparted by them is imparted by and not merely to their physical, active bodies. Otherwise there is no agency. I will assume, with everyone at the outset, that the world contains agency of the kind found in some of a person's comings and goings, movings and changing of things.

Agency is exhibited in more and in less sophisticated forms, that is, in any sophisticated, artful activity and in less complex, non-articulate physical activities. In both there appears to be more than mere causal efficacy imparted to the environment by a person. In sophisticated agency activities are organized, guided, purposive and purposeful comings and goings, movings and changes. And purpose is not absent in less sophisticated purposive activities of active creatures. So I shall argue in what follows. Now is the time for introducing the themes, topics, and issues to be considered, and the plan and purpose in them.

Sophisticated agency, as I dub it, is organized, guided, terminated, in accord with a person's view of things, his or her plans, interests, judgments, imaginings, desires, hopes, and *intentions*. Among the activities which instance such agency I would naturally list 'trying to find something,' 'testing a belief one has about something,' 'arranging things in a certain order,' 'carrying out a plan,' 'looking at and observing the changes in something,' 'telling someone something,' 'refusing to do something on order,' and many others from a variety of such activities. As this list suggests, there is no *obvious* association between so-called sophisticated intentional activities and so-called 'mental acts or activi-

ties,' nor a correlative, natural association between less fully intentional actions and so-called 'physical actions.' Every type of agency exhibits ability, capacity, skill, or, in short, power. Active power is required for any exercise of agency wherein a person initiates, guides, terminates actions, and succeeds (or fails) in changing things. In this book I shall map and survey the terrain in which agency plays its role.

The difference suggested above, between less and more sophisticated instances of agency, is crucial to producing an accurate map and a useful set of boundaries in an account of actions. Accordingly, major themes of the theory propounded here include the nature and variety of the abilities, capacities, proclivities, endowments, talents, and skills of agents, including prominently, their physical capacities and the stage of development of these capacities. Among the barely examined suppositions of our thinking is the idea that the more complex capacities, exhibited in sophisticated agency, are developments from and, thus, dependent upon the simpler, less sophisticated capacities of the kinds of agents we are. I accept this dogma, use it, rely on it. So part of what follows can be characterized as an exploration of the idea that fully intentional activities, involving sophisticated cognitive and conative attitudes and conditions, are only possible because of action capacities of persons in their physical and animal make-up. How this theme is orchestrated and whether it is successfully woven into a picture of the entire range of human activities is one of the burdens of the following chapters.

A second though less fully developed theme lies further in the background of the map I shall draw. It is introduced with the idea that it is activity itself which is the fundamental category, the genus, the *fundamentum,* the substrate of human action, in V.C. Aldrich's apt phrase the *field* of action, throughout the variety of human actions. If humans and other animals really are agents in the sense taken here, then any basic characterization of them features their native activity, in opposition to passivity and reactivity. Agents are decidedly not merely a kind of being *to* which efficacy is imparted. Given this theme, the question "why do they do anything at all rather than remain inactive?" is not the most natural question to raise concerning agents. Given this theme, it is more natural to ask why they do what they do rather than some other thing they can do, or would or might be doing. In short, my supposition here is taken from the fact of the ubiquity of the question "what is the animal, person, agent doing now?" To this question there is always some answer. To bring this theme fully into the foreground would require a metaphysical theory, a naturalistic account of such activity, and an 'energic' account of

causation, among other things. The full story is not to be met with in this book; but there are hints.

Somewhere Wittgenstein says that the concept of thought is a highly ramified concept, a motley, employed for a vast array of purposes, and so on. My third theme is that this holds true for the concepts required for an account of agency and is principally true of the concept of intention. Consider the diversity among the following. One may now have an intention to do some thing at some future time or an intention for the future. One may now have an intention to do some thing at a definite future time. One may have an intention in doing what one is currently undertaking, a further intention in doing what one is doing intentionally. One may have an intention in now acting, an intention with which one is currently acting. Further, there may be an intention in what one is doing in a sense which differs from both a further intention and an intention with which one is acting. A procedure, practice, mechanism can have an intention in it, as the intention in a breaking mechanism. One entertains, considers, formulates, reasons about, reasons from and to, accepts, and rejects intentions. Intentions are consistent or incompatible with other intentions and, perhaps, with beliefs. One expresses intentions in words and manifests them in deeds, both prior to and after the performances to which they are germane, as well as concurrently with actions. Intentions can be revised in the course of acting. Intentions upon which one acts guide behavior, express both means and ends, both goals and instrumental ways to goals, and express purposes. One can be doing on purpose something one had not intended to do. And, this is only a partial sketch of some of the ramifications of the concept of intention. The chapters which follow try to do justice to this variety without the total loss of a systematic and synoptic view of intentions.

The ideas expressed in these themes are hardly original. Nor could they be expected to be news if they do, in fact, characterize the perspective we take on our activities, undertakings, performances, and trials and successes. The problem is to understand them rightly, sanely, and fit them together with everything else we think about ourselves and the remainder of the world in which we act. There is no insufficiency of data for this undertaking. The problem, here, is to manage these data without being either overwhelmed by them or unappreciative of their extent. In presenting the view of this book I have tried to avoid the *eisitige Diät* of cases Wittgenstein warned against; a onesided set of data undernourishes theorizing about intentions, intending, action and agency.

It is not an arbitrary decision to begin with the so-called folk-psycholo-

gical or common sense view of human agency. Any attempted rational decision not to do so seems to call into play that very view. Further, how else might one begin the theoretical *enterprise* of understanding actions and agents in a realistic way? So the practices in which these beliefs and attitudes are present will be in focus in most of the following chapters. In addition to attending to the details among the resulting data, we need to manage the variety of cases with generalizations and classifications, designed to sharpen and highlight salient features. My procedure, then, shares with Kant's in moral theory the progression from ordinary knowledge of action and 'popular' rational psychology to an account of the underlying generalizations and, in turn, to critical theory of human agency. But to include and begin with common-sense views of agency is not to exclude other crucial sources of data. As I will insist throughout, human agents share a good deal of psychology of action with other animals, as an evolutionary and naturalist perspective will suggest. It is foolish *a priorism* to exclude any sources of data at the outset of theorizing.

The common-sense psychological view of agents and intention and the explanatory role of the last is contained, at least, in our linguistic practices with a family of words clustered around 'intend,' 'intended,' 'intending,' 'means,' 'action,' and others. I hasten to add, however, that not everything we need to know for our purposes can be learned from detailed considerations of 'what is said and thought' and the circumstances of so saying and thinking. Since it is clear that linguistic theory learns from such data, these data cannot be irrelevant to our undertaking. As well, psychological *theories,* motivation theories, conceptions of agents from other areas of human and animal science, comparative development studies, comparative psychology in general, and the account of the kind of physical organisms we are, are sources of relevant information, for all these bear on what we are and how we can be the kinds of agents we suppose ourselves to be.

In short, common understanding suggests that we are agents and initiators of changes, only some of which are fully intentional undertakings. However, but for the fact that not all such undertakings are fully intentional, with conscious cognitive and conative antecedents, none of them would be such sophisticated doings. These rational capacities of agents who are, as we think, both rational and animal, are developments from and dependent upon 'animal' abilities. The same common understanding, analyzed and dissected, presents us with a complex network of ideas and concepts, each with a function in our account of ourselves and others. So, if agency is not an illusion, we must also be the sorts of pro-

ducts of nature who can satisfy these ideas and concepts. For this reason the best available account of our physical structure and function and our evolutionary origins, that is, our natural base, must cohere with our view of intention and agency.

II

If anything is an agent and certain events are due to agency, then an agent must be such that it both can do and can not-do some, at least, of the things it does. And, its doings must be such that some, at least, of them both could be done and could be left undone. This is something like the notion of the freedom of indifference. It is, as well, a pervasive feature of our normal attitude about our own performances and the doings of others. We think ourselves both able to do and able to not-do the things we do as agents. Further, we recognize that the so-called freedom of indifference has a limited scope, in that not everything attributable to agents are things they could have avoided doing. We praise, blame and excuse actions; but we also exonerate and exculpate agents when we assess actions and find that their agents were constrained to act, in one or another of a variety of types of constraint. Conditions of diminished freedom are recognized as departures from the norm. Normally, agents are not constrained by 'outside forces'. Accordingly, we take it that our conception of an agent has these two features at the very least: (i) an agent's efficacy is imparted by and not (merely) to the agent, and (ii) as an agent, one is both able to do and able to refrain from doing the things which fall within the normal range of his or her agency. The attitudes which we thus take towards agents are enshrined in practical principles of morals and other conventions which we learn and teach.

In addition to the presumption of this sort of freedom or power to act, there is a further presumption of agency, *viz.*, that a normal agent has knowledge or reasonable beliefs about the nature of his or her actions. In various types of assessment of actions and agents, leading roles are played both by actual knowledge-like beliefs about the nature of our undertakings and by what any agent *should* think about his or her actions. We normally assume, then, both the capacity to do or not do a certain thing and the capacity to know what it is that we do.

These suppositions concerning freedom and knowledge are relative to the circumstances of agents in that the role of each will depend on various other conditions of the agent and the agent's circumstances. Such conditions include the age and development of the agent, the social and cultu-

ral setting of the agent, the agent's capacities and skills, and constraints and necessities in the setting of proposed and of executed performances. Presuppositions and circumstances of actions attributable to agents form a complex network of conditions. Obviously there are no easy articulations of rules for the application of or even the description of *all* such conditions. Our very practices make room for new and un-anticipated conditions which limit the normal suppositions concerning agency. Nonetheless, it is clear that there are established practices, perfectly applicable without considerations and thinking in normal circumstances, and there are recognizable departures from these practices of assessing acts and agents.

The foregoing description of the phenomenon of agency already reveals a host of concepts for theorizing. They yield such questions as: What is the nature of an agent's initiation of a change? What is the nature of the power or ability expressed in so-called freedom of indifference? What is it to attribute a doing to an agent, to another or to oneself? What are the circumstances of actions in contrast to actions themselves? What cognitive relations to circumstances and to actions must agents bear in their intentional actions? What is the nature of such relations? These are, of course, ways of stating some of the typical issues in the theory of action and agency. It remains possible that some or all of them are misbegotten. The final truth (if any) could, perhaps, still give the lie to the presuppositions of common practices. The possibility of universal mechanism or the possibility that the only events there are are ones which could not have been otherwise, or physical reductionism of an eliminative sort, or even the possibility of radical skepticism concerning the world and ourselves are possibilities which, if realized and confirmed, would arguably undermine our common beliefs and attitudes. Nevertheless, there is good reason for theorizing in spite of any ignorance we suffer on these matters. In the absence of an understanding of our common views, attitudes and practices, we will not understand the illusion of free agency, should it be an illusion. It is surely of interest to know what it is we are missing in case the world is altogether unfriendly to our attitudes and practices. A theory of intention and agency would provide understanding of 'real appearances,' however merely apparent they might be in the end.

This, too, is a Kantian theme in terms of which I understand the work in the following chapters. Nowhere do I undertake to fully justify the common practices and attitudes which characterize the common sense view of agents, actions, and intentions. Without the presumption of most of them, I should not know how to engage in an attempt to prove the

reality of these attitudes and practices. And it is sufficient justification if they can be seen to be coherent in themselves and consistent with what we think we know about the world in which we plan and act. This is demand enough.

By way of additional introduction let us begin to survey the kinds of data we will be concerned with in what follows.

III

For a variety of purposes, philosophers often look to forms of expression and their possible, plausible uses. For present purposes, some features of the nature of the required antecedents of intentional action can be detected thereby. Consider only the central notion of intending.

A standard form of question is "What is John's intention?" One standard looking answer form, "John is intending . . ." will appear odd, unless it is provided with a reading making it a way of expressing what "John intends to . . ." expresses familiarly. Thus, the present progressive of "to intend" can seem odd. Evidently, "to intend" differs from "to think" in this respect. As Wittgenstein reminded his readers, we can interrupt someone in thinking but not in intending.

What is to be made of these facts? What might be the interest in the fact that two important verbs, otherwise evidently similar in their forms and uses, differ in an important respect, and what might be of interest in the fact that a certain form is grammatically deviant? Of course, it is always possible that these are no more than accidents of language, unconnected with any important conceptual differences and similarities. At the outset of theorizing, however, I shall accept the dictum which requires that all data be taken seriously. It is an additional theoretical burden if a theory must explain some of its initial data as merely accidental, mere appearances. Thus, the present facts suggest that if someone is doing or has done something thoughtfully, thinking what he is doing or what he did, then there is a truth about him expressible by such sentences as "He is (was) thinking that he should do such and such," "He thought to himself 'I must now do . . .'," and similar sentences. But, by contrast, if someone is doing or did some thing intentionally, it is not required that there is a truth about him expressible by such sentences as "He is (was) intending that . . .," and similar sentences. Now, a theory which sustains this difference will have implications concerning the extent and character of the required antecedents of intentional activities. The data suggest (but do

no more than suggest) that an agent's doing some thing intentionally does not require that antecedently he 'is (was) intending' that thing.

Data of this sort can do no more than suggest pieces of a theory, since more than one theory will cohere with these sorts of data, and, since, in fact, the data are more often than not mixed in their suggestive importance. For instance, the present suggestion must be consistent with another matter of data, *viz.* if the question should arise whether someone did a thing intentionally, one good answer that he did will have the form "Yes, he intended to do it," and this form certainly appears to refer to an antecedent psychological or mental condition. The mixed data will be accommodated, in short, by a view which permits a distinction between those intentional doings which require and those which do not require that the agent, prior to acting, *had* an intention answering to his action. A completed theory will need to explain this difference and cater to all the distinctions needed to accommodate it

In pointing to the existence of intentional actions which were not preceded by articulate intentions which the agent *had,* these data also suggest an important parallel between verbal and nonverbal action. Most of human speech is spontaneous, unrehearsed, unstudied. Still, it is mostly significant, meaningful, secures reference and is largely apt, warranted, assertable, perhaps true. Similarily for human activities generally. It is mostly habitual, unrehearsed, unthoughtful, but apt and purposive. Unrehearsed speaking is typically meaningful speech; unrehearsed acting is typically intentional activity. The parallel and its implications for an account of intention and intentional action will be under examination in much of what follows in these chapters. But already it seems clear that no simple and no unilevel conception of intentional agency is going to accommodate the data we will survey.

More particularly, it already appears that the notions of intending, intention and intentional action do not form a neat conceptual fabric so that we could settle with a simple theory of intentional action. One such simplified and one-leveled theory might claim that (given the performance):

(1) An agent does something intentionally if and only if he/she was intending to do that thing.

(2) An agent was intending to do something if and only if he/she had an intention to do that thing.

(3) An agent had an intention to do some thing if and only if he/she was in a certain sort of conative-cognitive state (or stood in

such-and-such a relation to an appropriate sort of conative-cognitive object).

A theory which endorsed some form of (1), (2) and (3) would have a parallel in a theory of thinking, according to which an agent does something thoughtfully just in case the agent was thinking about what was done, and was thinking about what was done just in case certain thoughts occupied his mind, and this in turn is so just in case he was in a certain cognitive state or bore some appropriate relation to an appropriate sort of object, e.g., a sentence, proposition, or state-of-affairs, etc. Even if a basic theory of thoughts, thinking, and thoughtful action along these lines was acceptable, the corresponding theory of intention, intending, and intentional action will not be acceptable. No *neat* set of connections will accommodate even the sparse data already gathered, even at an unsophisticated level of description of those data. When we reflect on and theorize about our common concepts, we often expect a smooth surface and what we get in fact is a nest, an apparently tangled network of notions. (Wittgenstein reminded us of the tangled *(verworren)* texture of our concepts.) Accordingly, our theoretical target should be as much system and order as the facts will bear, when they are fully examined. The theorizing which follows will recognize the tension between the demands of theory – systematic ordering, simplicity, and range – and the tangled data.

The complex picture we will be trying to understand is also evident in the notion of intention itself. Consider. There seems to be a similarity between the propositional attitudes, such as believing, hoping and thinking that, and the practical attitude of intending. For instance, "John believes it's raining" appears to warrant the inference "John believes something" and the inference "There is something John believes." One might suppose that similar inferences are permitted in connection with "John intends to leave early." So, this might be taken to imply both "John intends something" and "There is something John intends." Now, traditional views about propositional attitudes such as believing hold that the verbs of propositional attitudes take objects and these objects are propositions. On such views, "John believes something" is the same as "There is a certain proposition and John believes it." A simplifying assumption might then be introduced, taking its cue from the already supposed similarity between propositional and practical attitudes. The assumption is that the verbs of practical attitudes also require objects of the same type as those of propositional attitudes. Just as one can be said,

suppose, to believe that a state-of-affairs obtains, one can be said to intend that a state-of-affairs obtains. Thus, on the resulting unified view of objects of attitudes, "John intends something" is the same as "There is a certain proposition and John intends it."

However, there are complications which suggest that such a simplified and unified view cannot be correct. For instance, the complement of 'intends' cannot be "that John will leave early," since John, in intending to leave early, might not intend that *John* leave early, where the last requires that John represent himself as John. Compare "The man in the corner in the garish green suit intends to leave early." The truth of this attribution evidently does not require that the man in question represent himself as one in a garish green suit. If a proposition is the syntactic object of the verb 'intends,' it is not one easily located even with the latitude afforded by indirect quotation.

Even if this consideration is not destructive of the view that the objects of 'intend' are propositions, other considerations, added to these, begin to tell against this unified account of the objects of the attitudes. For instance, the verbs of the practical attitudes (paradigmatically 'intends') *can* take simple infinitive constructions. *To leave early* is what John intends. One might suppose that "There is something John intends" is just "To leave early is something John intends." This suggests that generic acts are the objects of the practical attitude. While propositions and their kind might be required for the propositional attitudes, acts and their kind now seem required for the actional or practical attitudes. So the unifying view of objects becomes less plausible.

More complexity can be found. Suppose, what appears plausible, that if John intends to leave early and Bill believes that John will leave early, then what John intends and what Bill believes are the same. This supposition seems to support the unifying view of objects. Furthermore, intentions, like beliefs, appear to have implications. And they appear to be consistent with or inconsistent with both other intentions and with thoughts. How could this be the case if they are not the same sorts of things as thoughts, in the sense, now, of propositions? Once again, complexity appears at the outset in the unsophisticated account of data such as these.

If we add some thought experiments to these supposed facts, things become yet more complex. For example, it seems that there could be a language with expressive powers comparable to (perhaps fully comparable to) our languages but in which the psychological and practical attitudes had no explicit forms of expression. In the imagined case, speakers

and thinkers do not have such expressions as "I think . . ." and "I intend . . ." Their thoughts and intentions are expressed in 'free standing' declaratives which, *for us,* would be the complements of the verbs of attitudes. As an intermediate case one might conceive a language in which there was a system of voluntary, bodily movements which expressed the attitudes when made in conjunction with the verbal expression of 'free standing' declaratives. We, in fact, only sometimes employ the verbs of propositional attitudes and practical attitudes and decisions. We can express beliefs and intentions in 'free standing' declaratives, e.g., "It is raining" and "I shall leave early." Contexts, inflection and other indicators can make it clear that the speaker is expressing a belief (rather than a hope) or an intention (rather than a prediction). Nevertheless, a language lacking explicit cognitive and conative forms could be a language of agents and thinkers. Reflecting on these agents and their language suggests that the objects of the attitudes will have to carry the implicational consequences, the bases for generalizations, and so on, all on their own. Hence, it will be natural to think that they are propositional or sentential items. Opacity, then, only appears to arise, in part from the explicit presence of the attitudes, when opacity is interpreted so as to require implicational (and referential) inertness. The role of opacity, of course, introduces added theoretical burdens; at the very least it clouds the problem of the nature of the attitudes, since (i) it appears to run counter to the view that such objects carry some of the implicational structure, and (ii) nonetheless such objects, when not dressed in explicit forms, evidently do carry the implicational structure possessed by beliefs, intentions, hopes, fears, etc. A reasonably adequate theory must cater to (i) and (ii).

An account of the complements of 'intends' as fused expressions, e.g., to-leave-early-now, say, as the content of "John intends to leave early now," might account for some of the implicational structures we find. For instance, it might account for the inference from "John intends to leave" to "John intends something." The doctrine of 'fused whole contents,' however, runs into difficulties with such inferences as

(1) John intends to eat his artichoke.
(2) John has no evil intentions.

Thus,

(3) To eat one's artichoke is not an evil thing.

And,

(4) Intending to eat one's artichoke is not evil.

The inferences from (1) and (2) to (3) and to (4) appear to require that the complement contents of 'intends' are structured, have form, or even, logical form. Treating 'intends-to-eat-his-artichoke,' as an express property of John's does not account for the intuitive inferences, unless, of course, we adopt an extraordinarily, and hence unlearnable, number of 'meaning postulates' or 'meaning axioms' for such predicates.

IV

Theoretical requirements also arise from the content of current, relevant theories themselves. For example, any account of intentional agency will, in light of contemporary theorizing, have to be placed in the context of causal theorizing and deal with objections to the possibility of causal theories. No theory will be minimally satisfactory unless it comes to terms with the current issues over causal theories of practical reasoning and causal theories of intentional action. At present the central issue in the debate appears to concern the difficulty in providing a causal account of intentional actions which avoids the problem of wayward, internal causes. In addition to problems concerning the analysis of causation itself, the crucial problem for causal accounts of fully intentional action is that causes appear to be relatively indifferent to the manner, in full detail and specificity, of the causal route, from cause to effect. We are confident that a set of conditions and a certain event are causally responsible for a given effect (say, the brittleness of the glass and the trajectory of the stone, etc., are causally responsible for the breakage) without being in a position, always, to specify within narrow limits and with a fine-grained analysis, the determinate causal route in the series of events from cause and conditions to effect. It is sufficient (and perfectly reasonable) that our normal causal belief be backed by other beliefs about the physical structure or physical nature of the things and properties involved in a causal network, without a detailed knowledge of such structures. Beliefs about causal events are backed by beliefs that there are unknown causal laws, instantiated in physical structures at an appropriate level of description, and that these laws also support counterfactual claims about such causal events. A great deal of ignorance about details of causal processes is tolerable in this context. But in connection with explanations of intentional actions the degree of tolerable ignorance seems less. This arises from the fact that, at the level of data, we believe that no action can be fully intentional if it is only the causal upshot of antecedent cognitive

and conative states. In addition, the upshot must be produced in just the right manner if it is a fully intentional action. And, the upshot must in some way 'match' its antecedents. The problem is to explain this *manner* and *matching*.

The last bit of data is clear from the usual distinction between cases in which agents get what they want or desire by their own agency in contrast to cases in which their actions are fully intentional. Thus, if an agent wants something, believes that he can get it by X-ing, has the opportunity to X, and other conditions (if any are necessary) are satisfied, then if the agent X-es and his X-ing produces the desired effect, then the agent has got what he wanted (though, of course, he might not be satisfied by getting what he wanted). The effectiveness of his action is, at least normally, sufficient for the agent's having got what he desired. By contrast, an agent who intends to A and intends to A by C-ing, believing that he can A by C-ing, etc., need not have done intentionally what he intended in case his C-ing was causally effective in securing A. Fulfilling one's intentions appears to require more that what is required for satisfying one's desires. The last is close to being a case of causal effectiveness; the first appears to require both that the upshot is produced (causal effectiveness) and is produced in just the manner the agent intended. This requirement of fully intentional performances is the basis of the problem of inner wayward causal routes, since, evidently, a causal route from intending to undertaking action cannot secure the full intentionality of action merely by its effectiveness.

The same sort of issue arises in connection with accounts of practical reasoning. Roughly, as taken here, practical reasoning is reasoning *to* an intention where the intention is either a resolve to act now – a mobilized intention or intention in the ostensive mode – or an intention to act when and if one believes the conditions and time are right – a generic intention. While there are various accounts of the structure of such reasoning, almost everyone follows Aristotle (for a bit) in thinking that such reasoning is causally effective, either in producing action if the 'conclusion' is a present intending (or 'volition'), or in arranging or re-ordering the agent's dispositions and proclivities, if the 'conclusion' or decision is to do something at a future time and place. The question of the 'validity' of such reasoning, the role of choices as necessary premisses, the role of subjective assessments of probabilities and outcomes, and the assignments of utilities or values are all questions which arise, normally, within the context of theories of practical reasoning as causally effective processes. Actions are said to be 'reasonable,' 'deliberate,' 'fully intentional' in

light of such processes. (Not that the presence of such a process is a necessary condition for a reasonable or intentional action; it might not even be possible to construct a process of the appropriate sort, *post actu,* from the rational agent's antecedent conscious or unconscious cognitive and conative states.) But where such reasoning is present and action in fulfillment of it is pursued, there appears to be a 'gap' between reasoning and acting which purely causal effectiveness will not fill. The agent's behavior is fully intentional only if it is properly 'guided' by such reasoning. The problem is to fill-out the conditions for such *guidance.* In short, current theorizing imposes conditions on the success of further theorizing about intentional agency.

The causalist strain in recent theory of action should, I think, be balanced by an equally crucial condition on our account of agency. Perhaps it will be admitted that we remain ignorant concerning the causal story of a good deal of human action. The etiologies we tell are rough and ready generalizations. We suppose that we will come to know more and more about the causes of human behavior, i.e., why humans behave (in general) in the way they do. But the crucial reflection about this situation is that we also suppose that our present ignorance of the causes of the multiplicity and variety of types of human behavior is remedial. Thus, it is unlike what we take to be the unpredictability of this or that individual attitude, response, or action. *This* unpredictability is an essential feature of our picture of human beings. The unpredictability of individual agents in various circumstances is itself part of our attitude in interactions with others. Some humans can be trusted, counted upon, are truthful, do as they say, respond in typical ways in typical circumstances. But just as surely, others do not. Nor does a generally predictable agent always respond as one would predict, given the knowledge of the agent's past performances. Some behavior and some responses are such that they genuinely would not have been predicted. (Set to the side for now the idea of so-called *in principle predictability.)* Given the occurrence of a piece of unpredictable behavior, we can study it, study its causes, and we may come to understand its origins in that agent and in that agent's circumstances. Still, our attitude towards human behavior is qualitatively different from the attitude we take to, say, our machines. Trust in machines is not of the same quality as trust and lack of trust in persons. We predict when reliance upon a certain machine is no longer wise. The machine has been in use too long, its parts are no longer reliable, etc. The unpredictability of human behavior *is* the degree of reliance we invest in it. Humans give their word, and it is reliable or not. Humans find themselves in new

and unique situations in which what they will then do is not to be predicted. Previously reliable agents can and do behave in unpredictable ways, even in familiar circumstances. We trust some of them to do their best; towards others we take a guarded attitude. And we cannot always say what they will take their best to be in every circumstance.

Wittgenstein asks "what would it be like if human behavior were not unpredictable?" He also says that our attitude towards others is an attitude towards a *soul.* Wittgenstein's point is, perhaps, that others are not just "an open book;" we cannot read off their responses as we can the specifications of a machine from a schematic drawing, a wiring chart, a structural description and the like. With persons, their individuality, subjectivity, and inwardness is not, like a surface or a formula, open to view. So we naturally take a different attitude towards them, in contrast to the surface view we take of our constructed, well-understood, simple machines. There is no real mystery about this. It can be *put* somewhat mysteriously in "My attitude toward him is an attitude towards a *soul.*" The point of this remark is, I think, only to draw the contrast I have been just now discussing.

There are, of course, things which are not persons but which lack the transparency of some simpler machines. Very complex computers might be a case. And some humans – some innocents, children and others – might be more or less transparent to others. Animals, domesticated and well-trained, are often very reliable. One cannot trust a wild animal, in captivity or in the wild, in the same way. Not even an expert ethologist will have a fully reliable view of many such animals. Calculating, thinking, rationalizing persons with perspectives of their own can be very opaque and unpredictable, even to themselves. As well, some of them are "perfectly reliable and trustworthy." In either case, our attitude is not the one we take to an unfamiliar mechanism, whose workings we think can be more or less fully understood. Rather, by contrast, our attitude is trust or else lack of it, both in various degrees, and always in particular circumstances.

Any realistic theory of the nature and dynamics of human action must cater to these data. The subjectivity and perspectivity of agents must be given its due in any such theory.

V

An account of the nature of intentional actions will have to be placed in a larger theory of events, since intentional actions are events. At the same

time, no simple identification of intentional actions as events will be possible, since, it appears, actions are intentional only 'under certain descriptions.' Events as causal sequences are not barred from their roles as causes and effects by virtue of their (true) descriptions. No doubt the notion of a thing 'under a description' is not always clearly explained; but the point of the notion is clear. Suppose that if John decides to play the piano and does so fully intentionally, his act of playing is an intentional one. It also consists in some event or series of events. (Or, in his exemplifying properties; one's theory of events will determine what one says here.) On some views of this matter, the event which is his playing might also be the event of his causing the baby to awaken. Yet John, we may suppose, did not intentionally waken the baby and did not intentionally cause the baby to awaken. *If* events are individuated by their position in causal sequences, one might think that the event of the playing and the event which caused the baby to awaken are one and the same. For instance, on the view which identifies the antecedents of intentional actions, e.g., desires, beliefs, intentions, as causes of such actions and, as well, identifies intentional actions as events, it seems to follow that John's causing the awakening should count as intentional, since it has, by hypothesis, the same causes as his playing and is the self same event as his playing. But his playing was intentional and his causing the baby to awake was not intentional. Thus, theorists have suggested that the self same event can be intentional 'under one description,' e.g. "playing the piano" and not intentional 'under another description,' e.g., "causing the baby to awake." This motivates, in part, the notion of *things under a description.* Obviously, other theoretical moves are available, such as the denial of the account of actions as events presupposed in the above. But the simplicity of the identification of intentional action with events is also theoretically attractive.

It is also obvious that the difficulties just broached connect with a recent controversy concerning the individuation of action. The 'Multiplier' view rests on the idea that distinct action properties distinguish different actions. The 'Unifier' account rests on the notion that the same action can exemplify different and distinct action properties or action descriptions. Accordingly, within the discussion of the issue of action individuation and identity, it will be necessary to deal with the Multiplier-Unifier debate. Here is another case in which recent theorizing imposes theoretical requirements.

Additional controversy has arisen in connection with the notion of explanation of intentional action. The causalist position will, in one form or

another, insist that explanations of intentional action are a species of causal explanation, in some preferred version of such explanation. The anti-causalist position is usually associated with a more or less explicitly Humean view of causation. The anti-causalist insists that explanation of intentional action (or, at least, of rational action) is one or another form of non-causal account, inconsistent with a Humean, causal explanation. Thus, one version of the debate has taken the form of a question whether reasons are (or can be) causes of the action they rationalize. In the discussion which follows a causal (but not recognizably Humean) account of intentional action is presented. One version of the causalist vs. anti-causalist campaigns is simply avoided. Still, this controversy is sufficiently rampant that no theory of agency and intentional action can avoid it entirely.

VI

Ordinary forms of discourse, common beliefs, common attitudes and practices are not the only sources of relevant data. Theoretical constraints arise from outside what traditionally pass for philosophical theories. Results of scientific studies and current theorizing about motivation, character formation, language learning, physical functioning and structure are also clearly relevant to the views I shall defend. More particularly, studies of development are especially central to the themes discussed in Section I above and to the general perspective of the views in the following chapters. Development, training, learning are, indeed, an ingredient in the view I present. A simple formulation of this ingredient will be introduced with the aid of some remarks of Wittgenstein, in the first pair of chapters. Even more simply, I take it as a fact that when first learning, say, the names of colors, two features of sophisticated uses are not included: (1) 'looks,' 'appears,' 'seems' do not occur; (2) the person of the learner does not occur as a perceiving *subject.* Thus, names of colors are not learnt as subjective appearances of things. We can equally adequately describe this learning as (a) he learns to say 'red' (and the rest) on seeing something red, and (b) he learns to call 'red' (and the rest) what we too call 'red'. *Seeing* need not figure in an accurate description of this learning, though it too may. (Cf. *Zettel,* 421–426.

In what follows I will formulate, investigate and defend the hypothesis that the developmental-conceptual facts about perceptual language and thinking (in connection with color words and concepts, for instance) sug-

gests and supports a like parallel in connection with action language, practical thinking, and activity. For example, when first trained to do various things, e.g., fetch, hold, give back, arrange objects, and so on, two features of sophisticated action and practical thinking are not included: (1) 'intends,' 'intended,' '. . .in order to . . .' 'do . . .by . . .' etc., do not occur; (2) the person of the learner does not occur as agent of what is done. Primary purposive doings are not learnt as executed first-person intentions or as actions for the sake of further purposes. We can equally adequately describe this training as (a) he is trained to perform certain purposive activities, and (b) he is trained to do what we too can do for a purpose.

Additionally, the vocabulary of 'looks,' 'seems' and 'appears' is a new, further possibility in the language of perception and description. "The red visual impression is a new *concept.*" (*Zettel,* 423.) We teach 'It looks to me . . .' and 'It looks to him . . .' after the mastery of other parts of common descriptive vocabulary. Only then does a person occur as a perceiving subject, in the person of the speaker or in the person of the one to whom experience is attributed. Similarily, the vocabulary of 'meant to . . .' 'intended,' 'intends,' 'did . . .in order to . . .' and 'did . . .with the intention to . . .' is a new possibility, a new move in the 'language game' of action. We can say that the intended action is a new *concept.* Of course, these matters are here only introduced, not defended.

VII

The chapters which follow focus on all of the above and on additional issues. The notion that agents are a species of natively active centers of change whose actions can best be seen as falling on a continuum from voluntary and purposive to fully intentional, purpose*ful*, and elaborately sophisticated activities is introduced in Chapter Two. This 'activist perspective' is more fully developed in Chapter Three, where it is contrasted with the 'passivist' and excessively cognitivist presupposition of traditional theorizing about action and intentional activities. Chapter Four attempts to make good the claim that the concept of intention is a complex, varied instrument of our thinking (and so, our speaking). Thus, here we gather additional data which the subsequent theory must accommodate. Additional data are gathered in Chapter Five, where I study the notion of the types of expression of intention. In Chapter Six I begin to examine some of the structural properties of intentions, their

implicational ties and connections, and the variable 'scope' of the forms of intention. Since we are supposing that intentions are contents (of whatever type) of conative states, intentions as contents are examined in Chapter Seven. The account of such contents which seems to fit the theoretical requirements is compared with and evaluated in contrast to other possible accounts of intentions. Chapter Eight contains an account of the role of intentions and of intending (the content and the act) in explanations of intentional actions, as they fall, variably, on the continuum of such activities. An account of the role of causation in agency and a partial account of the specific kind of causation relevant to agency is examined in some detail in Chapter Nine. The theoretical powers of the theory of agency contained in these chapters is examined in Chapter Ten. An indication of its empirical base is present in Chapter Eleven and a short summary discussion is added in Chapter Twelve.

Here, then, are some detailed studies of issues in connection with action, agency, intention and intending. I claim that the resulting theory is coherent in itself, answers some crucial issues about the nature of our view of human agency, provides solutions to some of the currently leading issues in the philosophy of action, and is in harmony with what (I think) we know about ourselves as agents in the natural world. I suppose that a fully completed, wholly unified, and fully formal theory might have been attempted. But the state of our knowledge of all that is relevant does not encourage such a project. Perhaps studying some of the facts we think we encounter in our attitudes and practices as agents can be prolegomena to a simpler, more unified theory than the one in this book. I rather suspect, however, that much more will have to be examined and discovered before we can determine whether it is even possible for us to command a *synoptic* view of ourselves as agents, given that our activities are the objects of our own active inquiry. It is appropriate to conclude this introduction, as it began, with a reflection on yet another Kantian theme. If inquiry, theorizing and investigation are themselves, as we think, types of human agency of the most significant and sophisticated variety, then inquiry into agency itself had best first concern itself with its own preconditions. While modest, this is surely a proper study.

CHAPTER TWO

PASSIVITY AND ACTIVITY IN INTENTIONAL ACTION

> If in voluntary action properly so-called, the act must be foreseen, it follows that no creature not endowed with divinatory power can perform an act voluntarily for the first time.
>
> William James

I

Traditional views of the nature of intentional action typically agree that an action is intentional just in case it was intended and a certain connection exists between the intention and the action. However, there is typically disagreement as to (1) what an intention is, (2) what intending is, (3) what the required connection between intention and action is, and (4) what is the nature of the action itself. Thus, theories have been produced in response to (1) thru (4). An intention might be a first-person proposition containing an action verb in the future tense, or it might be a special sort of prediction about the agent featuring an adverbial modification (intentionally) of either the action verb or a sentence modifier, or perhaps an intention is a special sort of content, not either true or false, featuring a special tie between agent and action (a special form of the copula) or a special action predicate. Intending, in turn, might be a propositional attitude of a standard sort (knowing, believing, wishing, etc.), a special type of attitude with a special sort of content, or a disposition with a special sensitivity to perceptual information processes, or self-command, a self-directed imperative. Similarly, intending to do a thing here and now might be an act of willing, or a volitional event, or a complex of desire and belief, or a rationalizing pro-attitude, and so on. Theorists can be expected to disagree whether intending is a passive, receptive psychological state, necessarily a conscious state, a state or process non-observationally known by the agent, necessarily defined by its consequences, and so on. "Action" as referred to in (4) might be a bodily movement, an instance of an action property, the result or upshot of intending, a chain

of events in accord with an intention, and so on. Finally, the connection alluded to in the the traditional view might be identity or another sameness relation, causality, a special kind of causality, or a non-causal but rationalizing relation, and so on.

Various combinations of these views produce the variety of theories of the nature and structure of intentional action. Resulting theories are typically embedded in metaphysical accounts of mind and body, theories of the ultimate status of agents and their performances. Thus, combinations of answers to the questions introduced above and their embodiment in more general philosophies of mind can be expected to imply contrasting accounts of free action and responsibility, differing views on freedom-determinism issues, alternative accounts of practical reasoning and differing accounts of explanation and justification of human action. Finally, these differences will in turn pair with different views in the philosophy of the social sciences and psychology.

My purpose here is neither to canvas the many types of possible theories nor to examine their consequences. The purpose here is to formulate a framework or set of presuppositions which contrast with many if not all of the traditional views of intentional action. In particular, I will explore the consequences of rejecting, in whole or part, the traditional view that an action is intentional only in case it is preceded by an intention, only in case it is intend*ed*. In a restructured framework the category of intentional action is wider than the traditional category of intended action. Within the restructured framework I will discuss the issue I will call "the question of passivity and activity in intentional action." In a preliminary formulation the problem is this: if the antecedents of intentional action are themselves activities and actions of agents (in any of the alternative formulations of this condition), are not traditional theories in danger of an explanatory regress? And if the antecedents of intentional action are passive realizations (acts) of minds or agents, such states, events or conditions which can 'befall' agents, are not traditional theories in danger of losing the pre-theoretic distinction between what agents *do* intentionally and what happens to or in them?

II

Rather obviously there is a level of actions which are performed by agents with intentions in the guise of conscious decisions, deliberation yielding resolutions and decisions, and other types of practical reasoning

with practical conclusions. No doubt there are equally obvious reasons for focusing on such cases of action since we are interested in those actions which are, in some way, the upshot of reasoning in a way which makes them open to questions of justification. The concepts of reasons and justification direct our interest to this level of cases. But it is not obvious that all intentional actions are cases of this sort. An initial concern with intentional actions of a certain sort should not be taken as a reason for rejecting the possibility of actions at another level or from another category of intentional doings. Unfortunately, traditional theories have not questioned the correctness of taking the field of intentional action to be co-extensive with those cases of special interest from the point of view of justification. It remains possible that this is not a theoretically innocent point at which to begin accounts of human performance and capacity. Thus, in what follows I will question the view that the field of primary intentional action consists in performances with antecedents such as choices, occurrent conscious beliefs and desires, formulated purposes, conscious means-ends deliberation, explicit practical reasoning and the like. What alternative presupposition concerning basic cases of intentional action is available and plausible? What are the theoretical benefits of the alternative presupposition, the alternative framework of cases of intentional action?

The alternative framework is suggested by several kinds of consideration, considerations concerning development and learning, considerations concerning plausible conceptual dependencies, considerations of linguistic and grammatical data, and considerations of the conditions under which sophisticated directives are of use to agents. The benefits of the alternative framework can only become clear by reference to the theory it generates.

Much of received opinion and developmental psychology take it to be obvious that children are not first taught and do not first learn to perform deliberative actions and deliberate acts. They are not initially enlightened about and not initially informed as to their needs, ends and purposes and taught to reason and calculate as to possible and, then, suitable means in the form of behavior which is likely to satisfy them. When first taught to do this or that thing they are not taught to engage in reasoning of any kind. Practical, means-ends reasoning is a new twist on what they can do and on what they have been taught and trained to do. At the initial stages of development and learning, deliberative, deliberate, thoughtful, purposeful performances are not in the focus of the adult trainers and teachers. And they are beyond the focus of the developing learner. They

appear only after the development of natural capacities of voluntary movement and acquired modifications of such natural capacities. These initial action capacities are not taken to be non-voluntary where exercised. That is not the attitude of the adult trainer of children. Nor are they taken to be involuntary actions. That is not even a sensible attitude for the adult trainer to take. They are, then, of a kind with voluntary and purposive, if not yet deliberate and purposeful, action.

There is an important similarity between, on the one side, the learning of the descriptive object language and its vocabulary in contrast to the language of appearances and looks of things, and, on the other side, the learning of the doings of things and its vocabulary in constrast to the language of intention, purpose, ends and means. There is a significant parallel: just as it makes no sense to try to teach "looks (appears, seems, seems to me) φ" (for a wide range of descriptive "φ's") before the child has learnt ". . . is φ" or "φ", just so it makes no sense to try to teach "ψ in order to . . .(intends ψ, ψ for the purpose of . . . , . . . by ψ-ing, etc.)" before the child can ψ and has learnt rudimentary action descriptions containing "ψ". The child cannot understand and appreciate or think and use "intend to fetch the ball," "press the button in order to . . ." "set out to . . .first," "purposefully arrange the blocks" and similar things, until it has a grasp of and know-how capability of fetching things, pressing something, pressing one thing rather than another, pressing in serial order, opening something, arranging things and so on. It must first have these in its behavior repertoire. These doings are the first learnt voluntary actions, just as color terms and names of familiar objects are the first learnt descriptive words and object identifications.

The revised framework of intentional action takes a cue from these facts (or beliefs) about development, learning, teaching and conceptual pre-conditions. Two features are significant. First, the development of deliberative activities and deliberated actions presupposes a variety of non-deliberated and pre-deliberative but voluntary actions. Second, whatever the best analysis of the contents of deliberation – in its varied forms – *is*, it seems clear that deliberation and consequent intending are directed to or on such contents (proportional, sentential, proposition-like, or whatever). If so, a learning and developing agent must already have appropriate actional contents available in its thinking and reasoning for purposes of deliberation and subsequent deliberated and fully intentional actions. He will not have the capacity to deliberate as to whether, when, why to A, whether to A or to B, etc., unless he can grasp to some extent, what it is to do A, to do B. One way a child could (and likely does)

come to grasp these as actions it can perform is by having been taught or trained to perform them and then or subsequently taught to identify them verbally. The child comes to be able to think of things as its actions and, more or less at the same time, to speak of them via their names and descriptions. Thus the child has a repertoire of actions and action 'vocabulary' prior to a capacity for deliberation and intending. Its early actions *are* voluntary and, it will soon appear, have significant marks of the intentional.

Considerations of language also are cues suggesting the revised presuppositions of intentional action. "Do such-and-such more carefully!" is sensible and possibly appropriate only if the addressee can do the thing in question. "Did you do so and so deliberately?" presupposes that the addressee can do so and so and has an appreciation of the sense of "doing so and so." Further, more or less sophisticated pieces of advice, mandates, injunctions and prescriptions as to action to be taken require more or less articulated intentions on the part of one who *takes* them. This is why it can be a joke to *tell* a child to perform some sophisticated activity. More importantly, it also shows why such action-directing expressions (bits of advice, mandates, orders, etc.) are inappropriate in basic learning situations; the learner's performance capacities are not sufficiently developed so that it can make use of such directives. Hence, it is not appropriate to issue him or give him such directives. When forms of such action directives occur in *basic* learning situations, they do not have the role they play in connection with articulate and developed learners. In basic learning, directives of the sophisticated form only accompany training but do not produce it. It is the training itself, in *its* various forms, which produces the basic behavior repertoire; these are the initial voluntary actions upon which further training and eventually deliberation and intending vitally depend. Cases of initial voluntary actions (I shall call them the "primary intentional actions") are the basics in the field of intentional action.

The usefulness of practical advice, mandates, directives and the like rests on the priority of practical capacities in the form of primary intentional action. Forms of the traditional account are not consistent with this claim. On the traditional account, if an agent has the linguistic and pure conceptual capacity to 'grasp' a sentence issued as practical advice, then the agent can entertain an intention (a practical thought content) which the advice contains. Now, if to entertain an intention is to be oriented to action in accord with the practical content of that intention, then any agent with the requisite linguistic and conceptual capacity, who enter-

tains an intention, will acquire an orientation to action without other practical, learnt capacities. But this will evidently not normally be the case. Suppose (somewhat fantastically) that an agent has no capacity to ride bicycles or any other similar conveyance. Perhaps he is a Twinearthling, where such conveyances are completely unknown. Yet he does have impressive conceptual abilities. Accordingly, on some forms of the traditional account, he could acquire a *practical* thought (an intention) from the directive we would issue in "Adjust the curvature of your bicycle's path in proportion to the ratio of your unbalance over the square of your speed." But it seems plain that no practical thought will be forthcoming for the Twinearthling from such a directive, given the absence of any appropriate behavior repertoire. A Twinearthling will not acquire the practical thoughts and intentions from his conceptual capacities alone, for he hasn't the primary intentional action capacities which such intentions presuppose. His problem is that he has not got the required action repertoire of voluntary and primary intentional actions which he can guide by deliberating and intending. This, again, suggests the priority of voluntary and primary intentional actions to sophisticated, deliberate, purposeful actions, where the last are of the pattern suggested by traditional accounts.

It may help to make the present framework clear if we compare and contrast it with some views of Anscombe. In *Intention* Anscombe presents an account which appears open to two interpretations. The first interpretation is not clearly consistent with the revised framework. The second interpretation is consistent with the revised framework; however, on the second interpretation Anscombe's view in *Intention* is not clearly consistent. Let us examine these matters of interpretation before comparing the view in *Intention* with the traditional framework and with the revised framework.

By the first interpretation of the *Intention* view I mean to indicate the well-known view of Professor Anscombe that the 'criterion' of the intentionality of action includes the applicability of the question 'Why?', in a sense she explains, when it is demanded that first-person answers or else answers which yield first-person answers are available. On this view the category of intentional action only includes the actions of articulate, linguistically and conceptually advanced agents. Thus, on this view it would be a mistake to suppose that the basic intentional activities of developing learners are *properly intentional.*

On a second rendering of the *Intention* view, the 'scope' of the definition of intentional action in terms of the applicability of 'Why?' (in addi-

tion to the other conditions of the definition) includes the full range of action, of certain cases of animal behavior, of actions of children, learning agents, as well as sophisticated agents. Only those who ascribe intention to young children and animals must have the requisites for applying 'Why?' in the relevant sense.

If the view of *Intention* is one expressed by the second interpretation, there is no conflict with the present, proposed framework. However, there is a possible inconsistency in the *Intention* view on this interpretation. The problem arises in connection with the notion of *attribution* or *ascription* of intentions. It arises as follows.

Initially, for Anscombe the "area of intentional action" is the range of events in an individual's history to which the question 'Why?' has application i.e., has answers of a certain sort, primarily the expression of the person's reasons for acting. In addition, such events are known to their agent "not just because he observes them" (See *Intention,* paragraphs 16–19 for the central parts of her view. Note, in passing, that the condition of 'non-observational knowledge,' if spelt out in terms of 'the ability to say,' will not be satisfied by developing agents.) It is clear that, in the primary cases, the appropriate answers to the crucial question 'Why?' will have or else will yield a first person form. That is, while "Because he did . . ." "In order to . . ." "To . . ." and "With a view to . . ." and others are appropriate answer forms, so too are "I meant to . . ." and "I intended . . .". And answers of the first form will yield answers of the explicit first person form. If agent S answers "Why?" concerning his action C with "Because C was a way to D," this will be because S can also answer with "I C-ed intending to D" or some such first-person form. Too, it will legitimize others to attribute an intention to him and in doing so, to attribute some form of first-person reference to him. Thus, "S intended that *he* D by C-ing" uses 'he' so as to attribute to S something S would express by "I intended to D by C-ing". If spectators of the performances of another are warranted in attributing an intentional performance to him, then he must be warranted in expressing an appropriate first-person intention. The last is among the truth conditions for such attributions or ascriptions.

The last point requires emphasis. The attribution (in one sense) of intention in action or intentional action to another implies that the latter possesses first-person intention expression ability, given the criterion in terms of the language of 'Why?'. This is clear, too, from a consideration of the forms of attribution, "A intended to φ" and "A intended that *he* φ", when each is used to characterize an actual performance of A. The first,

"A intended to φ", is consistent with A's not φ-ing but ψ-ing instead and is the form of attribution often used to contrast actual performance with the agent's prior intent. In this use, prior intention is implied; hence, to attribute intention in this way is to attribute the capacity of prior intention and, therefore, the capacity for first-person intention expression or thoughts of the form "I intended to φ". The second form for intention attribution and the deep grammatical structure of such attribution, "A intended that *he* φ", straightforwardly implies that A intended something which he could have expressed by "I intend to φ", since this is the force of this use of '*he*'. (The points hold as well for attribution of intention for the future, e.g., "A intends to φ" and "A intends that *he* φ".) Hence, first-person self-reference is among the truth conditions for the attribution of intentions in the present sense of 'attribution.'

Now Anscombe allows attribution of intention to animals (and young children) but correctly denies them the requisite linguistic and cognitive capacities necessary for first-person self-reference (See *Intention,* p. 5 and p. 86). Thus, either her position is not consistent or, more reasonably, she employs 'attribution' and 'ascription' in two senses, one which has self-reference among its truth conditions and one which does not. In fact, there is evidence that this is what happens in the development of her view. Her explanation of our attribution of intention to animals is an explanation *by analogy* with attribution to agents capable of their own answers to the relevant question 'Why?'.

> Since I have defined intentional action in terms of language – the special question 'Why?' – it may seem surprising that I should introduce intention-dependent concepts with special reference to their application to animals, which have no language. Still, we certainly ascribe intention to animals. The reason is precisely that we describe what they do in a manner perfectly characteristic of the use of intention concepts . . .(85)

What is perfectly characteristic of the use of intention concepts is that events are intentional only under certain descriptions or only as described in certain ways, where the descriptions are primarily those derived from answers to the crucial question 'Why?'. Neither animals nor children have such answers of their own; we provide them. Evidently, then, the sense in which *they* have the intentions we attribute to them is not that in which sophisticated agents have the intentions they can aver in answer to 'Why?'. Hence, the second interpretation, discussed above, appears correct, on the proviso that two senses of 'attribute' in the general context of "attribution of intention" are present in Anscombe's view.

The revised framework which characterizes primary voluntary action

as intentional, though not the consequence of consideration, deliberation and intending, does suggest that the concept of primary intentional action is "formally independent" of the remaining concepts of intention e.g., the concepts of first-person expression of intention, further intention, and intentions with which an agent is doing what he is doing. Anscombe has argued that the supposed "formal independence" of intentional action from the remaining concepts robs the nature of intention of its definitive feature. For instance, she argues that under an hypothesis to the effect that we have no capacity to express future intentions and intentions with which, the intentionality of action would be no more than "a style-characteristic of observable human proceedings, with which is associated the question 'Why?'" (*Intention,* p. 30) This hypothesis is satisfied by developing agents in learning and training situations; the child's reaching for and grasping a particular object from among others is not the target of the question "Why?" which the child can answer. Yet it is not clear that his performances lack all important marks of intentionality. First, it is clear that *we* say of the child's performances that "He is picking out such-and-such an object" or "He is trying to get such-and-such an object", and we say these sorts of things only if we do not think of his movements as reflexes, involuntary. His behavior has the mark of intentionality in such situations as the following: if we move the object away, he will appear frustrated, not select another, nearby item; he will make further movements towards the moved object, respond 'happily' when he reaches it, and so on. This can be a game and a form of play. Here it does seem that "the proceedings-in-given description are what bears the stamp of intention" (*Intention,* p. 30). Nor will we be prepared to say that the child's behavior *is* his trying to move an eight ounce object of such-and-such material (or other such true descriptions). Thus the concept of intention seems to have a basis in our descriptions of the child's play, in his primary intentional activities, for what he is doing will not be equally well described, *as action,* under just any true description. Nor, in that case, is this merely a "style-characteristic" which lacks the feature of being intentional only under selected true descriptions. What is lacking is sophisticated intentional action requiring first-person responses to the question "Why?". Once again, the present account is not inconsistent with Anscombe's analysis in so far as the last is restricted to the field of sophisticated actions.

Anscombe rejects the traditional view, of course. She rejects an account of intending in any of a variety of ways one could characterize as "interior acts". More importantly, the traditional view implies that if A is

an intentional action, then A is (or was) the material or actual execution or realization of an antecedent, interior act of intending. Anscombe's account departs from this view, not by employing a notion of primary intentional actions, but by insisting only that descriptions of intentional actions "are *formally* descriptions of executed intentions" (*Intention,* p. 87). Likewise, the revised framework implies that primary intentional actions are neither the actual realizations of intentions nor are descriptions of primary intentional actions formally descriptions of the executions of first-person intentions. In this connection the notion of primary intentional actions is employed purely attributively or ascriptively. Nor can we seriously attribute the essentially intentional actions to developing children, learners or animals. That is, they will not be suitable subjects of attribution of such actions as "mailing a letter", "signaling", "selling something", "putting in a bid", and so on. (These are actions such that if they occur at all, they are intentional, and hence essentially intentional.) Children will be appropriate subjects of attribution of purposive actions such as those named by "fetching", "kicking", "taking hold of", "dropping", "picking up", "placing", and the like (Cf. Anscombe, *Intention,* p. 47). These are attributed on the basis of behavior and circumstances in the environment. Hence, the concept of primary intentional action is not *formally* dependent upon the concept of executed intention. In describing a child's primary intentional actions we do not suppose that it is so acting in order to realize its articulate purposes.

The revised framework appears behavioristic; this is as it should be. We need only remember that early behavior is not said to be intelligent, thoughtful, amenable to purely verbal directives in contrast to training and so on, until the child can have a view to doing something by doing some other thing it has previously learnt to do. A quasi-behavioristic model for these early performances is not a mechanistic one; we do not train and condition our machines.

In sum I shall assume that the revision of the traditional theory is *prima facie* reasonable and provisionally acceptable. How, then, does it illuminate the question of passivity and activity?

III

One theoretical motive for the traditional view, in many of its forms, is that the *activity* necessary for intentional action will be secured only by an antecedent activity of willing, intending, desiring which gives rise to in-

tending, or other such active states or processes. The thought is that the activity distinctive of intentional action – in contrast to other occurrences or events – could only be accounted for by appeal to something else which is *active*. Thus, the traditional account of intentional action faces the following dilemma. Either the antecedents of what it calls "intentional actions" are active (themselves actions) or passive (in one or another sense). If they are active, a regress is threatened. If they are passive, receptive, responses, sentiments, and the like, by what right do we identify their consequents as the agent's doing, his intentional action? If they are passive is not the agent a victim of what are called *his deeds?*

The revised framework for the area of intentional action suggests that the trouble with the traditional account lies in its presuppositions themselves. The new framework suggests that the field which is given for theorizing and analysis consist of, *at the outset, and ineliminably,* primary intentional, purposive actions of the learning and developing agent. Such action is not intentional by virtue of any separable, psychological antecedents. Behavior of this sort is but the exercise of the natural capacity and power of humans to move in their environment and manipulate parts of it. These are their basic doings; they are the sorts of creatures which have the power of movement. Those of their movements and non-movements which can serve purposes, even those not of their own devising – some natural and some their cultural and social inheritance – are their basic intentional actions. As they learn what they *are* doing and learn what they can do, they learn to further modify the environment in light of their successes and failures at achieving results, both natural and conventional ends. At some point they learn to say and think what it is they are (or have been) doing. The early entries in a list of what they learn they can do are what Wittgenstein would call "protophenomena" behind the concept of intending and the concept of fully, deliberate intentional action.

Accordingly, the present framework provides an answer to the problem of the passivity and activity in connection with primary intentional actions. These are not fully intentional actions in that they do not require cognitive, conscious antecedents. Hence, the question of the passivity and activity of such separable antecedents will not arise. The antecedents of primary intentional and voluntary actions are, rather, physical capacities, social and cultural settings, and training. These antecedents are situations and facts which surround the developing agent. They are not his psychological and conscious deliberations – in any of the various forms of deliberation.

The revised framework will certainly not, automatically, provide a

reply to the passive/active dilemma for fully intentional and intended action, however. It will, on the other hand, provide a demarcation among purposive actions, between those primary intentional actions for which the dilemma cannot arise and those sophisticated intentional, deliberate actions for which the dilemma appears crucial. (In Section V of this chapter, I shall speculate as to the further fate of the passivity/activity dilemma, given that there is such a demarcation within the field of intentional actions.) Having shown that the revised framework does embody a possible concept of intentional action and a concept of some plausible usefulness in the theory of intentional action, I wish to further develop an account of the concept of primary, intentional action. The concept can be traced to some suggestions put forward by Wittgenstein.

IV

In *Zettel* Wittgenstein remarks that "A child learns to walk, to crawl, to play. It does not learn to play voluntarily or involuntarily" (587). A traditional theorist, taking voluntary and intentional to be closely tied features, may well object that facts about learning (and training) to the contrary notwithstanding, there *must* be some feature of property in virtue of which activity is voluntary (or of the primary intentional sort). So in the same remark as above, Wittgenstein asks (rhetorically) "But what makes its [the child's] movements in play into voluntary movements? – What would it like if they were involuntary?"(587). I think Wittgenstein's thought here is that we will have to regard natural but early activities of children in accord with the framework I have been outlining. Learned movements and natural movements in play are, realistically, voluntary activities and intentional (but not intended) performances. Given the natural, physical capacities of children, they initially *engage in* purposive, but not yet purposeful activities. To imagine them to be involuntary, non-intentional, or purposeless is to imagine, not the basis for further learned responses with intrinsic purposes and ends, but rather some mechanical happenings of a kind of creature which cannot become an articulate (or even inarticulate) practical reasoner and agent. The applicability of the notions of teaching and training appear to depend on our belief that parts of the learner's initial repertoire of behavior and response to teaching and training is not involuntary movement, not purposeless. Such responses are part of what Wittgenstein has called the proto-phenomena upon which the concepts of intending and intention rest and from which they emerge *without consideration.* Thus, for in-

stance, an early game can involve the child's 'refusal' to return the toy – a 'refusal' which is a new response in the game of passing the toy back and forth between the child and adult. At some point the child will learn and come to appreciate that this new response has the significance of willfulness. In this and other ways children come to learn what they *have been doing* (A new phase in the language game). There is room for decision and practical consideration in the child's developing activities only after it has learnt and begun to appreciate what it has been doing. This is part of Wittgenstein's meaning when he writes as follows: "There is a particular interplay of movements, words, expression of face, as of manifestations (expressions) of reluctance or readiness, which are characteristic of the voluntary movements of a normal human being. If one calls a child, he does not come automatically; there is, e.g., the gesture 'I don't want to!' or coming cheerfully, the decision to come, running away with sighs of fear, the effects of being addressed, all the reactions of the game, the signs and the effects of consideration" (*Zettel,* 594).

If we were to suppose that some feature, in additional to normal responses and natural capacities and circumstances, must be present in order that the child's movements are voluntary and purposive, than we would be thinking of the child's initial movements as either involuntary or non-voluntary. But if we think of them as involuntary or non-voluntary, that is, really entertain the supposition that they are quite comparable to mechanical and predictable movements, then we would deal with them, react to them, in importantly different ways. We would draw different conclusions from the child's movements. In some cases we would naturally think, not "Here is a headstrong child", but "Here is an unusual mechanism". Not, "Do that again!" but "Perhaps that will happen again if I do *this*". Wittgenstein's point is expressed in "One draws quite different conclusions from an involuntary movement and from a voluntary: this *characterizes* voluntary movement" *(Zettel,* 599).

These data suggest that the natural physical capacities and their further developments are the bases for secondary actions which embody means-ends connections. Performances of the first sort are also means to other actions. The developing agent learns that such connections obtain. Learning these connections the maturing agent will be able to perform further actions which are further things he can do just by virtue of his basic capacities and their developments. The agent will learn to appreciate these connections, rely on them, and then to consider them and whether and when to employ them. At the outset a child will have to learn that the very thing he is or has been doing is also a secondary action; it is

also something with a purpose or further intention *in* it. Thus, walking, grasping, holding, moving things, arranging and changing things will, in the right situations, also be instances of coming (going) to a certain place, taking something from another, giving something to someone, putting something *here* or *there*, putting things in some order or arrangement, turning something on or off, showing someone something, playing various games with things. At a still more sophisticated level, exercises of basic capacities will also be instances of signaling, greeting, giving signs of the presence or absence of something, and the like. Thereby, the further things he can do by virtue of his controlled capacity for basic performances become further intentions with which he can do things in his basic repertoire. Here, then, are the proto-phenomena of *further intentions* and *intentions with which*. The available range of intentional actions would be narrow and sparse if there were no further intentions, ends, and purposes *in* what we learn to do or naturally, initially do. We learn the significance of voluntary movements and refrainings by virtue of acquiring the practical knowledge of means-ends connections, causal recipes, practical methods, conventional significance, and the consequences and results of what we have been doing in our basic, voluntary performances.

V

It should be clear that this restructuring of the field of primary intentional action, the field initially given for theorizing, presupposes that the agents of action are active 'by nature'. Movements, motion, activity, locomotion, altering the environment are the normal condition of agents. If so, the central question of action theory in line with this perspective will not be why agents act as they do, as if their natural condition and normal state was immobile passivity. The view of agents as immobile and passive appears to be at odds with evolutionary theory. Adaptability often requires mobility in the environment. And even perceptual capacity and memory are dependent upon the capacity for activity, mobility, controlled movements. (Some of the scientific evidence for this claim is surveyed below in Chapter Eleven. See especially the work of R. Held and A. Hein discussed in this chapter. As well, see R. L. Gregory's *Eye and Brain: the psychology of seeing*, 1966.) The central question is why do agents perform as they do, as if their normal condition is activity and the range of things they can do is far wider than what they actually do. The fundamental form of inquiries into actions will be: Why this action rather than those? The perspective of agents as active in their normal condition

implies the reality of the freedom of indifference, i.e., the ability or power to do *this* and the power to do *that* or another thing. Indeed, freedom of indifference is a presupposition of the applicability of concepts and descriptions of intentional action – both primary intentional action and more fully intentional action, since I have explained the concept of primary intentional action in terms of what an agent *can do* and what more it can do in virtue of its primary abilities. And since the range of things agents can do is far wider than what they actually do on given occasions, there will always be acts they could have but did not perform; there will be capacities they did not exercise.

Freedom of spontaneity, i.e., the freedom (liberty) to do something because one wants, becomes relevant to the explanation of action in that it provides one kind of answer to the question why an agent did what he did and, more generally, why agents typically do what they do. Given that an agent had the power to A and the power to not-A (in one or more of its forms - the freedom of indifference) one form of explanation why it did, say, not-A in the form of B, is that it did B because of its desires in the circumstances. This concept of desiring is designed to fix the connection between the circumstances of the agent and which thing the agent did from among the things the agent could have done in those circumstances. Note, importantly, that this does not imply that the explanation of why S did B requires S's desiring B or to do B; the explanatory role of the citation of desire does not demand, though it permits, that the object of desire is identical to the thing whose doing is explained by reference to the desire.

This preliminary discussion of "desire explanations" and, by implication, other explanations of actions by reference to psychological antecedents, bears on the dilemma of activity and passivity in at least two ways. First, if desires as explanatory are relegated to the role ascribed to them here, they are not required in explanations of primary intentional action. In primary intentional actions the activity of the agent – that he is *doing* something – does not derive from a psychological antecedent such as wanting or desiring something, since it does not derive from any antecedents other than his natural powers and developing capacities. And if at least *some* of the intentional actions more developed agents knowingly perform are merely the primary intentional actions themselves with their learnt significance and meaning, these more sophisticated performances are but exercises of the same powers and propensities. These more sophisticated performances require no antecedent events other that those of primary intentional actions. Part of the force of the 'dilemma' is avoided by the rejection of the scope of the traditional presuppositions.

Second, in still other sophisticated cases of intended actions, explanations by reference to the desires of the agent are not explanations of what the agent is doing intentionally. Desires are unlike sophisticated intentions in profoundly important ways. Where intentions specify, often, what an agent did or plans to do or is doing, desires often only explain why an agent does what he does in the manner in which he does it – why he does it readily, or does it against certain difficulties or why he does it so rapidly, gladly, happily, and so on. While intentions as plans *guide* our performances, desirings 'gear' performances. In such cases, citations of desires as explanations go to account for a manner, rather than a content of action. Hence, when in these cases desires are explanatory they only explain the manner of activity, not the activity itself. So *pure* desires, as psychological antecedents and causes, explain such things as why someone was willing to forego so much in doing what he did; it was because he so wanted to do it or to get what he thought it a means to getting. The agent's activity itself will have an account apart from pure desires. If so, it is not obvious that such active antecedents as desiring are required for explanations of what the agent is doing, intentionally.

VI

The view of agents as in the main passive cognizers, activated on occasion by desires which somehow trigger motions of their bodies, has had a profound influence on the problems treated in philosophy generally and in the philosophy of action particularly. Seldom, however, is the traditional view blatant. It is rather a background against which issues are raised and theories proposed. One might call it a perspective, i.e. something from which things are seen but not itself an object of attention. In the foregoing I have called attention to some features and some consequences of this perspective. It is now appropriate to examine one influential case of the existence of the traditional view of agents.

CHAPTER THREE

INTENDING, JUDGING, AND THE COGNITIVE MODEL

> . . . neither can the calculative faculty or what is called 'mind' be the cause of such movement; for as speculative [it] never thinks what is practicable, it never says anything about an object to be avoided or pursued . . .
>
> Aristotle

I

Donald Davidson's important paper 'Intending' contains an articulation of what Davidson calls "pure intending." (Citations throughout are to Davidson's 'Intending' in his *Essays on Actions and Events*, 1980). What does Davidson understand by "intending" and "pure intending"? We can gather the following features from his remarks.

(1) Pure intending may not be the result of deciding to do a thing or deciding to try to do a thing. Neither need it be the result of deliberation or of forming an intention. Thus, it appears, an agent intends, in this sense, if he merely has an intention to do a thing, without practical reasoning lying behind his intention, either explicitly or in some inferred, ideal *post actu* form attributed to him by an interpreter of his mental life. Pure intending need not be generated by contentual desires and beliefs.

(2) Pure intending need not generate action.

(3) Pure intending, unaccompanied by prior reasoning and consequent action is a condition sometimes mysteriously 'explained' by invoking episodes or attitudes "like willing, mysterious acts of the will, or kinds of causation foreign to science." (83)

(4) A plausible (but incomplete) account of acting with an intention will not have a parallel in the case of pure intending, since the plausible account invokes desires, beliefs, reasoning and acting, which need not accompany pure intending. But the accounts of acting with an intention and pure intending must be consonant.

(5) Although of little illumination, we can say that pure intending is "an action or at least something the agent does . . ." (89)

(6) But, pure intending ". . .is not generally the aftermath of . . ." a performance. (90)

(7) Intending is not believing that one will act. Nor is it believing not on the basis of observation that one will act. No doubt one who intends often does believe that he will do as he intends. However, one who intends to do a thing need not believe that he will succeed.

(8) "In the case of pure intending . . . the intention is simply an all-out judgment (not hypothetical or conditional and not a *prima facie* judgment that something is desirable). Forming an intention, deciding, choosing, and deliberating are various modes of arriving at the judgment, but it is possible to come to have such a judgment or attitude without any of these modes applying." (99) "To intend to perform an action is, on my account, to hold that it is desirable to perform an action of a certain sort in the light of what one believes is and will be the case." (101) ". . . a judgment that something I think I can do, and that I think I see my way clear to doing, a judgment that such an action is desirable not only for one or another reason but in light of all my reasons; a judgment like this is not a mere wish. It *is* an intention." (101)

(9) "There remains the question whether the sort of judgment to which I have referred, an all-out judgment, can be understood without appeal to the notion of intention or will." (101)

(10) ". . . intending and wanting belong to the same genus of pro attitudes expressed by value judgments. Wants, desires, principles, prejudices, felt duties, and obligations provide reasons for actions and intentions, and are expressed by *prima facie* judgments; intentions and the judgments that go with intentional actions are distinguished by their all-out or unconditional form. Pure intending constitutes a subclass of the all-out judgments, those directed to future actions of the agent, and made in the light of his beliefs." (102)

(11) Such all-out judgments take types of actions under one or another value aspect as their topics or objects. Intentional action with an intention here and now contains an intention of the remaining subclass; these are all-out judgments that this is a desirable thing to do. This difference reflects the fact that intending for the future refers to future rather than present acts. Pure intending is no more, nor less, mysterious than such all-out judgments.

Among Davidson's theoretical motivations for the features listed above is the advisability of avoiding an attitude other than some form of believing or judging, in order to account for the data concerning intentional action and intending. Pure conative attitudes, volitions or acts of will can appear mysterious in themselves; as part of the supposed etiology of intentional action and rational action, they can appear to require

notions of causation "foreign to science." T'would be best if they are unnecessary in accounts of intentional action as well as in accounts of pure intending. No doubt we understand the logic and semantics of belief better than the structure, 'truth' conditions, and reference needed for proposed conative conditions. And we have some grasp of the role of desire in explanation of behavior, both folk psychologically and scientifically. Pro-attitudes seem to us fit for ordinary causal accounts. "Wanting to see Jane at once and thinking her to be next door, Dick hurried next door" appears to express a causal account of Dick's strange behavior at midnight. No need here for acts of will energizing Dick's musculature; no need here for 'agent causation.'

Pure intending, however, appears to be a real phenomenon. Unlike Dick's intentional performance, produced by his pro-attitudes, Dick's intending itself need not issue in action. Nor need it result from Dick's cogitations about Jane and her whereabouts. If it is a genuine phenomenon of our psychology, it appears to be special, and hence, it appears to thwart Davidson's admirable motivatons.

We are not lost in obscure attitudes with peculiar features and extraordinary causal powers, however. Davidson's account is just that pure intending is no more than a judgment (a thinking that something is so). "To intend to perform an action is, on my account, to hold that it is desirable to perform an action of a certain sort . . ." (100) Dick intends to rush next door to see Jane. Dick judges that it is desirable that he rush next door to see Jane. He also judges that seeing Jane is desirable. What Dick thinks is that his rushing next door to see Jane is a good thing. Dick, of course, may have his reasons, good or bad. His intending to rush next door to see Jane is his all-out judgment, perhaps as a consequence of his reasons, directed to *his* future action, that his rushing next door to see Jane is the thing to do.

While we are, no doubt, presently ignorant of the causal mechanisms (if there are any) or of the causal law and its predicates which our explanation of Dick's behavior (if he acts) exhibits or exemplifies, still we believe that no 'mysterious' agencies are or will be required as we penetrate the physical underpinnings (or the psychological causes if different from physical underpinnings) of anyone's (and thus of Dick's) intentional performances. All-out judgments of desirability can be understood without unwanted appeal to the notions of intention or will, according to Davidson.

II

If we grant that judging and believing are more 'respectable' notions than intending and that the last can be explicated in terms of judging and believing, we still need to see whether all-out judgments will fit the theoretical bill. Davidson takes "judgment" in the sense of "believing" where believing need not be, though it may be, the result of considerations, deliberations, decisions, choices, selections, reasoning of some type or the result of other judgments. One often finds that one believes something, for instance, when asked one's opinion on a matter not previously considered or not the focus of one's previous thinking. In reply to a question, one *responds* and *reacts* with "I believe . . ." or simply with a free-standing sentence in the indicative or with assertive force. Pure intending is a kind of intending which permits the absence of antecedent rationalization. Thus, both judgment and intending will be explicated independently of antecedent rationalization and decision. Nor does pure intending require consenquent action. Just as a response that one believes a thing to be so need not figure in any subsequent behavior or thinking, pure intending need be no more than either an unrealized disposition or a temporary but causally and rationally inert condition. Accodingly, both all-out judgments and pure intendings will have be explained by citing their 'intrinsic' features.

The relevant 'intrinsic' features, of course, will be features of expressions of beliefs, judgments or intentions. For instance, Davidson is not engaged in a psychoneurological or an introspectionist study of judgments and other attitudes. Concentration on the logical or quasi-logical forms of expressions and their roles or possible roles in reasoning (or the absence of any such roles) will indicate differences and similarities among the attitudes. Nonetheless, it is equally clear that the attitude under study is, on Davidson's view, a cognitive one. It is, roughly, holding true a propositon or judging a thing or sort of thing to have a feature of some sort. In particular, it is judging something of a sort to have the property of being desirable in some regard or other.

Consider, then, the possible 'objects' of expressions of intentions. In connection with intentional actions, actual rationalized performances, the subject or topic of a judgment will typically be the action itself. The concrete, particular undertaking of eating the pie in front of one is what is judged to be desirable here and now; its desirability in some regard (e.g., sweetness) is judged to override any other considerations the agent has. For instance, he might have promised someone that he'd not eat sweets

today. But he now holds, judges, that this pie is so delectable and his promise was so silly and ill-advised, that, accordingly, he now will have this pie. Here is a specimen case of intentional action where the action itself is the topic of one's intention.

In the case of intention for the future, of which pure intending is the genus, there is no concrete, present undertaking, no action, to serve as topic of the desirability judgment or intention. Thus, it appears reasonable to suppose that pure intending or the expression of pure intention takes the values of the variables of quantified sentences as subjects. What one judges to be desirable in this case is some (maybe even any) act of pie eating. This will not do, however, since (so-to-speak) unintended models will satisfy these judgments. Some pie eating episode might satisfy the predicate ". . . is an eating of a forbidden piece of fattening food" or, even, ". . . is an eating of a piece of poisoned pie." Let us suppose that these are not what one rationally judges is desirable. Furthermore, an action in accord with one's intention for the future can always be reasonable, for the agent, in case the intention was reasonable in light of what the agent thought and desired, and he has not altered his desires, and no evidence requires him to alter his beliefs. But if an intention for the future is a general judgment (universal or particular form of quantification) it can also be unreasonable if it issues in action which satisfies his intention, for he would 'intend', say, to eat any or every piece of sweet pie he encountered, never mind any other feature of the pie in question. The problem is to find something other than concrete actions, on the one side, and values of quantified variables of general sentences, on the other side, as the topics of pure intending.

Davidson's suggested way out of this situation "is to make a firm distinction between the kind of judgment that corresponds to a desire like wanting to eat something sweet and the kind of judgment that can be the conclusion of a piece of practical reasoning – that can correspond to an intentional action." (97) The last are singular judgments about the action undertaken. They have a demonstrative component – what one is undertaking – and feature it as desirable to do, given what else the agent believes at the time. Pure intending or judging that some sort of action is desirable in some respect does not permit the logical detachment characteristic of quantified sentences or judgments. The judgment that a certain sort of action is desirable is, rather, a judgment that a type of action is desirable in a certain respect. Thus, it is a *prima facie* desirability judgment. An all-out judgment, by contrast, is a further judgment that no other considerations are sufficient to outweigh the desirable feature of

the type of action under consideration. Davidson concludes, therefore, that "the judgment that corresponds to . . .the action cannot . . .be a *prima facie* judgment; it must be an all-out or unconditional judgment which, if we were to express it in words, would have a form like 'This action is desirable'." (98) However, Davidson does not identify all-out judgments expressive of intentions *for the future* with demonstrative judgments. Were he to do so, he could not insist, as he does, that in pure intending the intention "simply is an all-out judgment." (99) The all-out judgment expressive of an intention for the future is a judgment in the form of an "interim report," *viz.*, "given what I now know and believe, here is my judgment of what kind of action is desirable." (100) This is not a *prima facie* judgment expressive of the conditional desirability of a type of action. Rather, it "assumes, but does not contain a reference to, a certain view of the future." (100) It is conditioned by the agent's beliefs, but it is not the conditional belief or conditional judgment that an action of a certain sort is desirable if certain conditions hold in the future. In short, intentions are what is expressed by either demonstrative judgments that this action is desirable or by all-out judgments that a certain sort of action is desirable, where the last presupposes the agent's assessment of conditions but does not itself express those assessments and conditions. Judgments which express intention are, thus, forms of belief about present action or about types of action, featured as desirabe, in all-out unconditional forms.

III

With this gloss on the view, we must ask if all-out judgments that an action of a certain sort is desirable will indeed bear the role of an intention for the future. To test this consider a case in which present actions are undertaken (or dispositions acquired) with further intentions for the future. Davidson's own example of a person acting so as to provide for the future welfare of his heirs is such a case. The agent undertakes actions *now* with a view to providing for the future; or, the agent acquires a disposition, realized in his habits of investments and savings of income, say, with a view to the welfare of his heirs. Intending for the future can be a reason for the agent's present actions and dispositions. Thus, why does he now set aside a certain portion of each salary check? He intends thereby to provide for his heirs. He does it, intending to provide for them. These last facts go to explain his present habits and actions; they give one of his reasons for such behavior. He judges that ceratin present actions and habits are desirable because he judges that it is desirable that his

heirs are provided for and he thinks that his present performances are likely to contribute to their welfare in the future.

Schematically, agent A has a reason for φ-ing now in that he judges that ψ-ing in the future is desirable and that φ-ing is a reasonable means to ψ. His reason will typically be expressed in something with the form "ψ-ing is desirable," where this is an all-out judgment.

Consider, then, the role of propositions of the form "φ is desirable." An agent might certainly consider, reason about, accept that something is desirable and not intend to pursue it as something he will try to bring about by his action. Although the desirability of the future welfare of his children would be an understandable reason for a man's present and future action, its actual role as his reason depends on something like his use of it in rationalizing his actions and its role in actually guiding his performances. There does seem to be this difference: (a) The agent judges that something is desirable but this belief does not function as his reason for action, and (b) The agent judges that something is desirable and his belief does become his reason for action, where his action (a) and (b) is, plausibly, the same. He writes the same will, saves and invests the same amounts, and so on in each case, and in each case, he judges that providing for his heirs is a good and desirable thing. Still, in one case he intends *by his action* to provide for his heirs and in the other, he does not have *this* as his reason for the actions he takes. In case (a) the agent knows that his children, say, will probably receive the benefits of his savings and investments but that is not his reason for his behavior.

The view under testing will at this point invoke the difference between *prima facie* and all-out judgments. The claim will be that in case (a) the agent thinks that the future financial welfare of his heirs is *prima facie* a desirable outcome, while in case (b) he judges, all-out, that it is desirable that they be provided for. (In both cases, of course, he judges that a sort of action, not a concrete undertaking, is desirable.) This distinction need not appear *ad hoc*, concocted for the purposes at hand without independent motivation or reason, since, it will be insisted, in case (a) the agent could be reasoning that if various contingencies actually transpire, another outcome would be preferable. For instance, he might reason that if his heirs should attain wealth on their own, inheritance of his wealth will not be needed and he would prefer its use elsewhere; then, if he decides that his heirs are not going to be in need, and makes other provisions for his wealth, he will not have contravened his previous plan and intention *and* he will not have changed his mind, adopted a new and contrary intention. In (a) it was not his intent to provide for them, though

he knew that his actions might well have that desirable consequence. In case (b), any other intended outcome in the absence of a change of mind or heart on the agent's part will be a contravention of his previous plan and intent. Purposeful action contrary to a previous judgment of the *prima facie* desirability of an outcome can be perfectly reasonable; this is part of what the *prima facie* feature of the judgment means. Purposive action contrary to an all-out judgment that something is desirable is not reasonable on the part of the agent so long as he has not rejected the desirability judgment which provided his reason for acting.

But is this reply adequate to account for the difference in the data, the difference cases of type (a) and (b) provide ? The (a) and (b) cases and their differences seem to show that judging a thing to be desirable and possible is not sufficient for the role of a further intention with which one acts, when such a further intention is the agent's reason (intending) for acting as he does. The issue, then, is whether or not this role difference is merely the difference between judging that something is desirable *prima facie* and judging, all-out, that something is desirable. First, it is certainly not obvious that the distinction between *prima facie* and all-out desirability judgments can carry this explanatory burden. For instance, one might judge that a certain outcome is the most desirable probable upshot of a contemplated course of action, decide to pursue that course of action, actually undertake the course of action in question, and yet not have adopted the further intention whose content contains that desirable outcome. It is unclear how the *prima facie*/all-out distinction will deal with judgments that something is the most desirable among a set of alternative outcomes. There is a range of differences among desirability judgments with no clear, non-question begging way to apply the *prima facie*/all-out distinction within the judgments in this range. Second, one might reply to this difficulty with the claim that a comparative desirability judgment is an all-out judgment is case (i) it is the judgment that something is the most desirable outcome among the contemplated alternative, and (ii) it functions as a further intention. But this would not be an available claim for the Davidson-style view, since it is a requirement of such a view that intending is to be explicated in terms of all-out desirability judgments, without appeal to such notions as intending itself. Condition (ii) violates this requirement. On the reductive view of Davidson, it is not supposed to be possible to identify all-out desirability judgments by appeal to a pre-theoretic notion of intending. In the absence of such an identification, the data suggest, though they do not prove, that one may well accept that ψ-ing is the most desirable probable outcome of his φ-ing and

undertake to φ but not with the further intention (and reason) to ψ thereby. The all-out desirability of ψ, which he accepts, is not *thereby* his reason for φ-ing, unless part of the meaning of "all-out" is that cognitions with this feature thereby play actual roles in reasoning and thinking which can lead to rational action – the "all-out" feature thus riding conceptual piggyback on the notion of intending and further intention. (Cf. Chapter Four for further considerations on desiring as necessary or sufficient conditions for intending.)

What, then, could be lacking, if all-out judgments or beliefs that something is desirable are not sufficient for further intentions in acting and, thus, for reasons for acting? I wish to consider this crucial question in a round about way, *via* a speculative digression in conceptual history. My claim will be that this question seems so crucial and central to us because of certain natural, yet possibly gratuitous, presuppositions.

IV

Why, suppose we ask, does desire and desiring ('springs of action') and the desirable (things featured as good in some measure of value) play such a large role in recent accounts of rational and intentional action? What explains the intuitiveness often felt in connection with the view that desires and desirability judgments are primary in accounts of why persons act as they do? Could it be that it is only in virtue of an implicit presupposition about agents and agency that this intuitiveness is felt so strongly?

These questions appear, rhetorically, to suggest that the intuitiveness just remarked on is a symptom of another implicit theory and the thought that this theory expresses a deep truth about agents and agency. This suggestion can be reinforced by noticing some few facts aboout agency. It is not often noted that there is a difference, no less obvious than the importance of desire, between the conditions under which agents satisfy the complex predicate ". . . got what was desired by acting and the complex predicate ". . . achieved what further he/she intended by acting." The first complex desire-action predicate is satisfied by an agent who (i) had a desire for X, (ii) performed some action, Y, (iii) produced X by the action Y, and, perhaps, (iv) produced no other consequences sufficiently contrary to his/her set of desires at the time of acting. But the complex intention-action predicate is satisfied, roughly in case (v) the agent intended to Z by Y-ing, (vi) intentionally Y-ed, (vii) brought about Z by means of Y, and (viii) Z was produced by Y in a manner or *via* a connec-

tion which was, roughly, anticipated and counted on in intentionally Y-ing. The last condition, (viii), has no counterpart in connection with getting by one's action what one desired. Within rather wide limits, any means of getting what one wanted will do; within rather narrower and more specific limits, only an anticipated means of achieving one's further purposes in acting will count as fulfilling one's purpose intentionally. The difference between the truth or satisfaction conditions of these two predicates or relations is, of course, the basis for the so-called problem of wayword causal routes in the causal account of intentional action. The presence of this difference and the problems it raises should suggest to us that an account of intending in terms of desires and desire-judgments alone might face significant problems, in spite of the 'intuitiveness' of appeals to desires in explanation of action.

Furthermore, appeals to desires in action explanations play various roles. "Because she wanted φ" or "Because φ-ing is desirable" (as expressed by her) might well explain why she did what she did, given her other preferences and beliefs about means to φ. But such citation can also explain, not *why* she did just what she did but, rather, why she did what she did in the manner she did it. "Why did she respond to the speaker's question so quickly? She was the first to respond, Why?" asks for an account of the way or manner in which she acted. These questions do not focus on why she responded, rather than leaving the meeting, say. "She replied quickly because she so wanted to attract the attention of the speaker" might explain why she was quick to respond. Another sort of case: "Why did he give up his vacation so as to finish his report to the Society?" asks why he was prepared to forego some desirable state or consequence (attributed to him in the query). The answer can be given by citing the strength of his desire to further the purposes of the Society, for instance. In these two sorts of cases citations of desire go to explain (a) a manner of action rather than the action itself and (b) the choice of one of a set of alternatives each of which is desirable.

Finally, one might develop some suspicion concerning the role often assigned to desire and desirability judgments by noticing that for many cases the notion of desire will have to be very general and very remote from states which serve the felt needs and wants of the agent. "Why did she visit her dying relative?" might get answered, truly, by "She felt an obligation to do so; she really would have preferred not to; she did not want to but thought she ought." If one insists that 'feeling obliged' and 'thinking one ought' are forms of desiring, it is clear that they differ from other clear cases, such as desiring to eat, sleep, desiring fame or fortune,

wanting to be successful, to be noticed, to be happy, healthy, and the like. The role assigned to desire and desirability in the general sense of these notions is close to accepting 'psychological hedonism', a general theory of motivation. That theory is either not obviously correct or else is so general as to be uninformative. Again, we have some reason for a counter suggestion to the traditional role assigned to desire in the account of action. Of course none of these data are sufficient to refute the traditional view; they are, however, counter suggestive. This is enough to motivate a further inquiry into the presuppositions of the crucial and central role typically assigned to desiring.

In fact, a model of human agency stands behind and supports the role traditionally assigned to desiring in theories of human agency. Here is one more version of that model.

It has been remarked, with boring regularity, that our theoretical forebearers e.g., Descartes, Hume, Kant and others, adopted a 'perceptual model' of knowledge. Less repetitiously, it is worth noticing that they also left us a cognitivist model of persons or agents. Like many of their philosophical and theological ancestors, any *motive, appetitive*, or *active* function was thought to be irrational, or non-rational, or rational only if controlled by right cognition. Cognition, too, is passive. Our primary (or sole) contact with things outside the mind is passive reception of 'features' or 'feature-like' impressions of things "without the mind". The most accurate, significant judgments, for Descartes for instance, require willing assent; but, the accuracy of judgments depends on the will's total domination by cognition. The tradition of Hume insists that activity, conation, will, is but an impression, passively received like all impressions, when, as Hume says, we are aware of "knowingly giving rise to a bodily movement." Finally, we could gloss Kant's central issue in connection with the possibility of pure practical reason as the question *how can* willing be dominated by reason and cognition alone and agents moved to act in disregard of appetite? This tradition conceives of agents as passive, cognitive receptivity machines, capable of change and movement when sufficiently agitated by externally produced impulses. Desire, in one or another conception, was a natural phenomenon to play the role of activator, 'spring of action'; nothing less seemed sufficient to move passive, ponderous, cognizing machines such as human agents. It is, then, plausible to suppose that the role assigned to desire is a consequence of the cognitivist conception of mind in modern philosophy, in its predominating tradition.

The cognitivist model implies the native passivity of agents, since

agents are 'essentially' minds, receivers of impressions, information, input – let the category be as ancient or current as you like. Thus, when considering the fact that agents seem to act, seem to initiate changes and movements, seem to "give rise to voluntary movements" and so on, the natural, irresistable question is why do agents act at all rather than remain in their 'natural place' of passive receptivity? The theoretical need for 'springs of action' has been thus created by the cognitivist model and its implications. The notion of what the agent desires now seems to fit the theoretical need produced by the cognitivist-passivist model or presupposition.

No doubt there are reasons for and causes of the traditional model, in its various forms. Epistemological issues in modern philosophy are among the causes. The impressive physical theories, the idea of causality and natural, scientific lawfulness which philosophers thought these theories required, and the antiteleology and the reductionist tendencies of some early modern philosophy are among the reasons one can so easily think that what must be avoided, in any reasonable account of action and agency, is "kinds of causation foreign to science." (Davidson, 83)

What I wish to suggest is that in spite of the reasoned cultural cause of the traditional model, the domination of the model is an historical contingency. *It* contingently structured the questions, problems, and possible solutions in philosophy of action and agency. To see this more dramatically let us engage in some highly speculative, creative reconstruction of large parts of modern theorizing. Suppose, if you can, that Descartes had been, say, a pragmatist, and that Hume, say, much more of a pragmatist that he is in the *Enquiry*, and that Kant and the tradition had inherited problems from an activist, conativist, evolution influenced, non-epistemologically dominated past.

As I intend this peculiar hypothesis, the idea is that the conative attitude rather than the cognitive dominated, was more 'basis', captured the core notion about humans, and the contents or 'propositional objects' of this attitude were not primarily the bearers of truth-values; rather, their primary features were effective ways and means or ineffective ways and means of achieving natural and conventional ends. Proposition functions, ". . . is φ," were not more basic or important in mental representations than 'practical-actional functions,' represented by ". . . to φ." Propositions, *that is a Q,* were not more central attitude contents than were the results of introducing personal constants into 'practical-actional functions,' producing structures such as "Helen to Q." Propositional modalities, 'Possibly (a is Q),' were no more fundamental that the practical modalities, such as 'Permissibly (Helen to Q),' 'Imperatively (Helen

to Q),' 'Advisably (Helen to Q),' and 'Irresponsibly (Helen to non-Q)' and so on. Perhaps the cognitive query 'Is it the case that P?' addressed to John, would be naturally expressed by, and its role filled by, 'Please (John to find out whether P),' and the assertion that P, addressed to John, would be role identical to 'Advisably (John to accept that P),' and so on. Propositional logic, predication, syntatic and semantic inference would not have had counterparts in practical inference, actional predication and corresponding syntatic and semantic structures and metalinguistic rules. What *we* call infinitives of action verbs might have been no less basic (more basic?) than propositional predicates and functions. ". . . to Q" would be a basic form for connecting properties and things, and some of the properties thus tied to individuals would be some of *our* cognitive verbs.

An underlying metaphysical view which might have accompanied this fanciful theoretical past is, indeed, a scientifically plausible picture. Evolutionary theory seems to demand that we and our embodying structures have been selected for our ability as adequate believers. Psychological structures have been selected to put us in reliable contact with 'the external environment.' From this perspective it is not possible that we (or our components) are subject to massive error in beliefs. From one evolutionary perspective, the epistemological question should not be "How is it that our beliefs are ever reliable indicators of the environment?" Rather, it should be "How is it that some error and unreliability is possible and, sometimes, actual?" Perhaps it has not been noticed that this question is analogous to the fundamental question Descartes' view of things required him to face. Descartes' Creator, of course, selected him, his mental structure and incarnate properties, so that the former would be a very reliable indicator of things which threatened the latter. The reliability of judgments of bodily pain as indicators of environmental conditions was selected for in Creation. The very same structures, bodily and mental, which produce this reliability are the structures activated in perception of the external environment generally. Yet, according to Descartes, by theoretical necessity the external world is other than it appears in perception. Hence, the question becomes "How is the presence of such error in perception to be explained?" Descartes' answer is, roughly, that the Creator could not find a better way, combining considerations of speed of information transfer, economy of structure and the like, a better way, that is, to select for the one virtue without providing for the other disability. But this is not, in general, a kind of move open to our evolutionary perspective. An evolutionary advantage would not be an advantage if it was gained by the selection of structures and powers which also produced

great disadvantages in the long run. An evolved world can be better than the best possible world in that in the former there can be fewer necessary evils. If cognitive capacities are the result of such marvelous selection, and they pretty evidently are, then massive error of belief is impossible. Epistemology which focuses on the problem of skepticism, turns on the question how reliable belief is possible, and worries how the isolated individual mind can get in contact with the ambient world operates with a model of knowing subjects remote from an evolutionary picture of agents.

Now, reliable belief and the selected structures which account for it are not isolatable from conative, active, exploratory, locomotive capacities and the structures which account for them. In philosophy of mind and action this has been recognized, though not fully exploited, in talk of the belief-desire-intention-action complex. The efficiency and effectiveness of any one of these is connected with the like virtues of the remaining. We have reason to think that we and other species could not have the reliable beliefs we have if we did not enjoy the powers of directed (desire and intention directed) action. Nor would our actions be usefully directed if we did not have reliable beliefs about the environment in which we act, about the progress in the world of our serial undertakings (through perceptual feed-back mechanism), and about the terminal success or failure of such directed actions. Capacity for successful action is surely as much a part of our actual evolutionary background as is the tendency to reliable belief. The suggests that in actual fact, as well as fanciful conceptual history, conation is no less native than reliable cognition. Conscious trees, so to speak, could never come to have beliefs and cognitions of which active, exploring, directed agents are capable.

In short, I am suggesting not only that the cognitivist, passive model of human agents is an historically contingent legacy, but that it is at odds with the very evolutionary models we have now. Hence, the problems, theoretical issues and requirements, which depend on and get their urgency from the cognitivist model should be replaced, where possible, by those associated with an activist model. If we perform this shift of models, what then can we say of the problems in the philosophy of action with which we began?

One such problem was to understand the role assigned to desire and desirability characterizations in the explanation of action. Various pieces of evidence suggested that it is a mistake to assign this role to desire and all-out (cognitive) judgments that some thing is desirable. The differences between the predicates or relations ". . . got what was desired by acting" and ". . . did what he further intended" was only one such piece

of evidence. Much of the evidence fits the activist perspective recently introduced. The requirement of a 'spring of action' which moves otherwise passive agents is relativised to particular circumstances and no longer a requirement for all actions undertaken. Thus, there is no theoretical motive for reducing the second of the predicates or relations to the first, no need, that is, to require that further intentions are reasons for or causes of actions only because of the motive power supplied by desire. The two predicates can be given separate readings, on the activist model; but a unified, desire reading is suggested by the cognitivist model.

The main practical question about action shifts, on the activist model, away from the ubiquitous, but simple question form "Why did agent A do X?" to the more complex "Why did agent A do X rather than any of the other undertakings open to A?" This shift occurs because the passivist presupposition requires an explanation of why agents do anything at all in contrast to their normal passive condition, whereas the present perspective assumes a large repertoire of action capacity on the part of the agent and a propensity to normally realize parts of that capacity. On the last model, the central practical question has the form "Why did A do X rather than Y, Z, W, . . . etc.?" In short, the activist model assumes the reality of the freedom of indifference, *viz.*, that agents are capable of and can perform a multitude of actions in any given circumstances unless certain special incapacitating conditions exist in those circumstances. On this view, agents are sets of powers and structures of dispositions to change the environment. Thus, the agent is expected to be active, realizing some of its action propensities in accord with its directing conations or intentions in normal circumstances. *Intending* is now crucial since it is a directing and guiding disposition, realized on occasion in conscious intending.

V

The Davidson style account of intending will be labeled, henceforward, the "cognitive-desire view." In concluding, I wish to test this model against the phenomenon of akrasia or weakness of will.

Akratic action is supposed to present a puzzle because it is thought of as intentional or voluntary action contrary to the agent's decided reasons and intending. Furthermore, akratic action is not a mere conceptual possiblity; there *are* instances of such intentional actions. Puzzles arise over how to describe and, then, how to explain such actions. Of course, one response to such a puzzle might be to deny the conceptual possibility of such cases. Since I take there to be akratic acts, I shall not consider this response. Further, the problem is not merely to provide a conceptually

coherent description of the possibility of weakness of will. In addition, so far as this is open to view, one must speculate about and evaluate plausible accounts of how such events occur. Both parts of the puzzle will present problems for the cognitive-desire view.

First, on the cognitive-desire view, in its usual causal form, an act is intentional just in case it was the causal upshot of the right sort of cognitive-desire antecedents and was produced in the 'right' way. An akratic act is one produced in the manner of intentional actions, since akratic acts are intentional. But, by hypothesis, the akratic action is contrary to purported – manifest and attributed – antecedent cognitive-desires. This fact suggests that cases of akratic action are cases of 'akratic desire', that is, cases in which desires are sufficient to produce appropriate action but fail to do so. If so, akratic action is evidently paradoxical in the extreme.

Second, the cognitive-desire view does not seem resourceful enough to account for akratic actions, even if they are possible on this view. That is, on this view there do not appear to be any resources available for counting such acts as both voluntary (or intentional) and 'weak-willed.' If they are the result of 'mechanical breakdown' in the agent, they are not voluntary or intentional. If they are the upshot of the cognitive-desire complex, they are not akratic. This result might be said to be due to my unsympathetic rendering or overly narrow account of the cognitive-desire view. It might be noted, for instance, that nothing in the view precludes a distinction between *levels* of cognitive states and antecedents, some of which levels are (or some level of which is) 'essentially' action producing, some others engaged only at a ratiocinative level. Thus, when the agent is 'of two minds' or 'at two distinct levels' with respect to an actual performance, akratic action is possible and is produced by the essentially action producing level of cognitive-desire, which is contrary to the ratio and conscious level. Meta-level desires might be invoked, so that the weak-willed agent both desires A and desires B, but his desire that his desire for A be satisfied is stronger than his meta-desire concerning B (though the last does give him a reason for acting on his desire for B.) But even if some version of a two-level view is not inconsistent with the essentials of the cognitive-desire model, this addition to the view creates a difficulty or two for it. It is, first, unclear what the higher level, conscious states are, since they 'float free' of action and action guiding roles (on the view). Second, the distinction among levels of psychological states, some essentially action producing and others not, pushes the cognitive-desire view closer to the alternative, activist model. It was a central feature of the last that basic conative states are essentially productive of performances;

cognitive, reflective states are reactions to actual and possible action scenarios. Attempts to incorporate some such feature into the cognitive-desire view blunts the distinctive character of it. Furthermore, it appears to be motivated in an *ad hoc* manner, just by virtue of the issues over akrasia. And, to the extent that this is more than a description of akrasia, i.e., more than an attempt to show that akrasia is possible, the view can appear to warrant the use of a notion of causation or productive agency which is, according to Davidson, 'foreign to science.'

Does the alternative, activist model account for akrasia? Certainly not automatically. It does provide for the notion of native voluntary and intentional (though not intend*ed* action of agents. It also provides for the sophisticated conation of intend*ed* actions and further intentions for action. Akratic action is characterized as intentional, in the sense of primary intentional action, even when there are further sophisticated intentions and reasons for the agent to act differently. And the activist account does not warrant a notion of causation 'foreign to science' (whatever this slogan means) since it leaves it an open question how to explain, causally, productive basic intending. (Akrasia itself will be examined further in Chapter Ten.)

In summary, we have noted that (1) indending and all-out (cognitive) desire judgments can play various, differing roles as reasons for action, (2) that the reason giving, causal, 'springs of action' role assigned to desires rests on a model of agents which, when it is generalized to all intentional action, is not itself forced on us, (3) one alternative model helps explain the differences there are between predicates or relations such as ". . . got what was desired by acting" and ". . . did what was further intended by acting," (4) the alternative model assigns a plausible role to desiring and desire judgments in explanations of the *strength* of action and reasons for the manner of action, and (5) the alternative model *seems* promising as a basis for an account of akrasia, while the cognitive-desire account appears to fail this test. 'Pure intending' will divide, on the alternative, non-reductive view, into basic conative states of active agents, and into the more sophisticated practical thinkings which are the further intentions with which self-conscious agents act. Intending is not desiring because desiring is not or need not be sufficiently 'cognitive'. Intending is not judging because intending is not or need not be *that* 'cognitive.' Intendings produce the active, directed change of which agents are capable. This notion of *producing* need not require an idea of causation 'foreign to science.' If our data are substantial, the idea(s) of causation involved in agency had better be accommodated by 'science' if they are not present there already.

CHAPTER FOUR

FURTHER AND FUTURE INTENTIONS

> A man whom I cannot deny, may oblige me to use persuasion to another, which, at the same time I am speaking, I may wish may not prevail on him. In this case, it is plain the will and desire run counter.
>
> John Locke

I

Much of what is to follow runs counter to a widespread rumor that once a person has performed a given bodily movement he has finished acting, the rest being up to nature, God, the friendly surroundings of the world and so on. This is a sort of physicalist version of an older version of a volition view of agency. According to the last, once a person has performed a given willing (or else once a volition has occurred) the agent is finished acting. If the agent sees or otherwise becomes aware that this acting on his part is not having the desired consequences, he performs new bodily movements (or willings, etc.) and hopes for the best. He keeps trying by his movements (or willings, etc.) and either the world conforms to his desires or it doesn't. This view is of a piece with the 'conscious tree' or 'inept brain' picture of, say, perceptual belief. He keeps 'recording' the passing scenes, and his records either agree with 'external reality' or they don't. Both are bad pictures of human agency and experience. As indicated in Chapter Two, both are rejected or drastically modified in the present account. Some further modifications can be introduced through an examination of data concerning further intentions and, then, intentions for the future.

Agents voluntarily and intentionally produce changes and alterations in the world in which they act, move, think and plan. Some to the things they do have their purposes *in* those doings themselves. Such actions are identifiable through their purposes. The further intentions with which such things are done are intrinsic to these undertakings themselves. Many such actions are forms of activity. An example of an action which is an activity with an intrinsic further intention is practicing scales on a

musical instrument, Playing the scales *is* practicing, in suitable circumstances. One may undertake to practice by playing scales with yet a further purpose, say, to earn praise from one's mentor. "He practices as much as he does in order to gain praise." "He plays the scales over and over in order to practice" sounds like a pleonasm because much playing of scales *is* practicing, in normal circumstances. (I shall omit the condition 'in normal circumstances' below; it is to be understood unless expressly denied.) A further intention "of the first rank" is an intrinsic purpose (practicing, as above); a further intention of any higher rank is a purpose with which an intrinsic further intention is pursued (to gain praise, as above) or else a higher rank further intention $_{(n+1)}$ is a further intention with which the further intention $_{(n)}$, itself a higher rank further intention, is pursued. Thus, one can play the scales with the further intention to gain praise and intend thereby to further some additional purpose, e.g., to please one's parents, believing they will be pleased by the praise lavished by the mentor. "Self-seeking" further intentions normally terminate in further intentions to gain some personal value, such as self-esteem and the like. Egoism is, in part, the view that every higher order further intention has this terminal higher order intention. Further intentions which are not intrinsic to the activities (i.e., not further intentions of the first rank) which have them I will dub "extrinsic further intentions." Not all higher rank further intentions are extrinsic, since some higher order further intentions have intrinsic further intentions of their own. Saying a certain thing in certain circumstances may involve the intrinsic further intention to warn someone of a perceived danger; it may be one's job and duty to issue such warnings, so that warning others of a perceived danger is undertaken with the further purpose of doing one's job properly. This purpose, in turn, may have intrinsically, an increase in salary or some sort of bonus attached. An intrinsic connection can obtain because of an institutional or a contractual arrangement.

II

By virtue of what facts or conditions is it true that an agent has a further intention (of either sort) in doing what he does intentionally? This form of question suggests that whenever it is true that an agent did something further or something else purposively in doing what he did, it will have been true that he *had*, before or while doing what he did, a further intention or purpose. It does seem typically true that one does something further in doing what one does intentionally; if it follows that one *had* a

further purpose, it then seems that the psychology of the agent is in danger of becoming overcrowded with purposes or, alternatively, that there are far fewer purposive acts than we are supposing there to be.

What then is it to *have* a further purpose? A further intention? This invites the question what is a further purpose, further intention? The short answer to the last question is that it is to act with knowledge of (or belief about) means-ends connections. "To improve as a flautist" names a purpose that some agents sometimes have. To have a particular purpose is not the same as to *want* the result described by the name or description of that purpose, although it is normally the case that the agent desires the result of his further intention. Wanting to achieve some end is neither a sufficient nor a necessary condition for having a further intention in acting, even where one wants to perform actions in furtherance of one's further purpose. A normally informed and rational agent may not want to bring about some end, Y, by X-ing and perform act X, and Y may be secured by means of one's X-ing; yet one may not have wanted to Y. Nevertheless, if Y is an intrinsic further intention it will be true that the agent intended Y (or to Y). One might intend to follow a certain order, want to follow orders and do what was ordered, though not want (and even want not) to produce the obvious consequences of the action required by the order. An officer may be ordered to send his men into battle, intend to and obey his order, but not want his troops deployed as they will be deployed by carrying out his order. Yet he intends them to be deployed as he was ordered to deploy them (against his better judgment). Wanting is not a necessary condition for further intending.

Similarly, one may do something, hoping it is a means to one's further purpose, want to achieve one's further intention, and yet achieve such a purpose only inintentionally (or non-intentionally). If one's action only accidently produces results which match what one wants and produces them in some way one had not anticipated or intended, then while one welcomes the happy upshot, it is not *fully* intentional. One might want the results of one's actions and intend them but achieve them in an unintended and accidental manner. Wanting to achieve a further purpose is not sufficient for its being intentional, in case it was achieved in some quite unintended manner.

To *have* a further intention in doing one thing intentionally is to have an action in prospect which is an end of one's intentional activity. In acting, to *have* a further intention is to have an action in prospect it is to be not unaware or not only later aware of an action in prospect. To play the flute with the purpose of improving is to play with improvement as a

prospect. To go somewhere *now* with the purpose of returning early is to go with the prospect of returning early.

There are various sources of prospective actions. To have decided, or to have chosen, or to desire in accord with one's knowledge, or simply to be normally knowledgeable with respect to means-ends connections are all such sources. Now it may be that theorists will want to restrict *having* a further intention to some one or some few of these sources. There would be no danger in such a restriction so long as it is recognized that in that case one can act with a further intention without '*having*' (chosen, say) that purpose. This seems overly restrictive, since it requres the qualification just mentioned.

Accordingly, the kinds of consideration involved in one's having a further intention are psychological and nonpsychological. That is, it may be by virtue of a choice or decision in their psychologically full-bodied senses that one has a further intention or it may be in virtue of one's circumstances and one's standard knowledge of means-ends relationships that one has a further intention. In the last, most prevalent sort of case, an agent need not have decided or chosen a given purpose in order to be acting with that purpose. He may simply know that what he is doing intentionally is a means, in these circumstances, to a given end; in these circumstances he is doing what he is doing intentionally with the further intention defined by that end.

It is only in the full-bodied uses of such notions as choice and decision that psychological states are relevant to *having* intentions or intending. For it can be true that an agent did what he did by choice just in that he could have done otherwise. Refraining was clearly open to him. And he did not consciously select the course of action in question. Similarly for decisions; having considered a course of action in light, say, of a certain problem, an agent may pursue a particular course of action. He is said to have decided on that course of action even though there was no specific conscious stage in his considerations which is identifiable as his decision to follow that line of action. Obviously, the full-blown use of these notions, for cases of specific conscious acts, is also sometimes engaged.

The distinction between further intentions of the two sorts delineated in Section I appears to correlate with the different conditions under which one has a further intention. Further intentions (of the first rank) which are fulfilled (to some extent or degree, at least) just insofar as an action is performed with such an intrinsic further intention are also those further intentions which one has, typically, just insofar as he has the normal range of knowledge of means-ends connections and intentionally

performs acts that count as means. Further intentions that require the presence of various sorts of external conditions for their fulfillment, once appropriate primary actions have been performed, seem typically to be ones an agent has by virtue of a more nearly fully conscious act or deliberative process. For if an agent doesn't know or believe (as part of his normal practical knowledge and skill) that his acts further a particular purpose, then it will not be his further intention in acting unless he comes to see such a connection or possible connection through deliberation, calculation, decision, and so on.

III

The attribution to others of further intentions of either of the main types discussed above is based on verbal and non-verbal behavior of the putative agent and on the nature of the circumstances of the action under study; and attributions of further intentions are made in conjunction with our view of the agent's beliefs and desires. In *Motive and Intention* (1972) Roy Lawrence discusses three instances of what he calls "intention invocations." They are (1) the claim that something must have been intended by what was done (intentionally), (2) the claim that it is possible that a certain thing was intended by what was done, and (3) the claim that a specific thing, rather than other things, *was* intended by or in what was done. What Lawrence demonstrates for historical narratives which invoke further intentions holds as well for ordinary accounts of human proceedings. Where further things are done in the activity of normal, knowledgeable agents, the further things done and the purpose are one, in the same sense in which, if I obey, then what I did and what was ordered are one. (Cf. *Zettel*, paras. 290–292) If an inference to a further intention is required, it is an inference from the primary thing done and the objective practices in operation, to the character of the further action. Standard, objective practices are in operation if (i) such practices exist and provide means for possible ends, and (ii) the agent in question is more or less competent in these practices. If the practice is not standard, objective and known, the inference is to the claim that something further must have been going on, though we as yet do not know what it was. We can speculate, however, since we can supply possible practices and possible means-ends relations which the practices could embody. If we can find no end for which the action under study could be a means, we may simply find the case, so far, inexplicable. Without the agent's desirability characterization of an end of the second sort (extrinsic purposes) we may not

have surveyed the proper sort of item in searching for the further intention in what was done, for the second sort of further intention is an end or purpose typically *via* some desirability characterization on the part of the actor or some characterization according to which it is the agent's purpose. If no such inferences seem successful, the possibility that there was no further purpose is open. We might be tempted to expand "characterization" in these last remarks and explain it as "*desirability* characterization by the agent." I do not think this would be correct if it is taken to require the agent's wanting the result in question. Explained as "desired or purposed," the notion of "the agent's characterization" will not be misleading; however, it will seem circular as an explanation. This, too, is among the motives of those theorists who propose that *wanting in the wide sense* is a condition of intending; otherwise, it looks as if we must say that adopting a purpose is such a condition, and that is circular! I have suggested that the realistic way out of this circle is to understand purpose through practical knowledge, skill, and acts with intrinsic further intentions; these just are among what is learnt in socialization, maturation, and culturation. Extrinsic purposes, *via* deliberation, invention, are creations of our making which presuppose our learnt capacities with actions whose means we learn in whatever way we do this.

If an inference to a further purpose conflicts with the agent's avowal (or disavowal) of a further intention, it by no means follows that his word is overriding. (Cf. Anscombe, *Intention*, paras. 25, 27). Lawrence presents a useful example of this. A Renaissance painting of Maffei had been assumed to portray Salome. Erwin Panofsky showed in *Studies in Iconology: Themes in the Art of the Renaissance* (1939) that the Francesco Maffei painting depicted Judith with a charger rather than Salome with a sword. His evidence consisted in facts about the subject of the painting, the existence of prior paintings depicting Judith with a charger, the nonexistence of any types such as Salome with a sword, etc. His case is made by these facts. Suppose, however, that evidence seemed to be found to the effect that Maffei declared of his painting that it was Salome. Maffei's word on the matter is not decisive. It is not privileged in and of itself, since it is inconsistent with the *facts* that Panofsky marshalled for his conclusion. As Lawrence observes, this ". . . is why the fact that the chance of there having been such an announcement (by Maffei) sincerely made, is very small really does counter a protest based on the undeniable *possibility* of Maffei's having made the announcement. For the very conditions which make improbable such an occurrence make it unreasonable in these circumstances to treat the sincerity of the speaker as a reliable sign of the truth of what he said." (p. 106)

Further intentions of the first rank are there, to be seen in the actions which are means to them. To the extent that circumstances are clear and unambiguous and the means-end connections are standard, in actual practice there need be no doubt as to an attribution of a further intention. Obviously, there are as well those cases in which circumstances are not clear or means are not standard for certain ends. Here, then, room for speculation, hypothesis, and evaluation (and perhaps, in the end, skepticism) is provided. But if skepticism were everywhere possible, we simply would not have the notion of purpose or further intention we clearly seem to employ. Enquiry concerning further intentions exempts certain purposes from doubt; an agent *does* have the purpose his action aims at in *these* circumstances.

The account of evidence of further intentions has so far relied on the notion of practical knowledge of standard means-ends relations in normal circumstances. While these features require exphasis, we need to turn to an issue that so far has been avoided, for surely there are ends and purposes that lie outside the tidy reach of normalcy. And surely agents pursue such ends in their intentional actions. Are there any limits on the intelligibility of further intentions? The crucial issue of evidence for nonstandard further intentions arises in connection with unfulfilled intentions. Given that an agent was intentionally performing some primary activity or other and given that he was not satisfied with the result of his activity, what *was* the action he had in prospect? He may simply tell us; we may or may not find his account believable, initially or in the long run. Without access to the word of the agent, we rely on behavioral and circumstantial evidence. What was the action in prospect? this is the same question as what did the agent think, believe, hope, etc., would be the upshot or result he would produce by means of his primary action. What he believes may be true, false, self-deceptive, mythological, wildly irrational, and so on. If it is true, then he has a new or an as yet unrecognized but available means for achieving some end or other. While his further intention may not be fulfilled, because of a failure in the supporting necessary conditions, his primary action may still embody a new piece of practical knowledge. He has a new handle on the external world, which, conditions permitting, can be verified by other agents.

If an agent's belief about a way of achieving a certain purpose is false, then *he* will not have succeeded in securing his further intention, even though he had that end in prospect and even though, in some cases, events which would fulfill his further intentions might occur. For others to understand his action, they must come to understand his false belief that the primary thing he did was a means to the end in question. This

formula covers at least two types of cases. The means-end belief may be false because *that* is not a way to secure this end, though another means would do. Or, the belief may be false because there is no way in which that end (squaring the circle, say) can be produced. Yet, some of our ancestors prayed to make it rain and Hobbes did what he sometimes did with a view to squaring the circle.

Wildly irrational cases can also be understood and verified by evidence. Sartre's Daniel says that he deliberately does the opposite of what he wants. One may hold an irrational belief that by doing A he will bring about not-A; his further intention in A-ing may then be not A-ing. Obviously no such means-end connection will provide a handle on the world of objects; it may, however, provide a recipe for some psychological states. This is a factual issue for the wise, the insightful or the psychoanalysts to discover. Irrational further intentions, like false beliefs concerning means-ends connections, may be due either to irrationality concerning the means to certain intelligible ends or to the irrationality of the end.

Evidence of further intention, in the last few kinds of cases, depends on our capacity to reconstruct structures of belief not our own, given that we do not participate in the 'irrationality.' In order to know what the evidence for a non-standard, because irrational, purpose would be, we must supply the surroundings for that sort of action. It can have a place, so to speak, by virtue of its surroundings, and prominent surroundings are the agent's beliefs. If we know that he believes that he can control the growth of his garden by his prayer, we can seek evidence of his further purpose in praying as he does.

Having stressed these points, one must hastily add that the 'objectivity' of our attributions of further intentions of the extrinsic sort should not be overemphasized. It seems clear that in actual practice alternative 'objective' attributions are available. Persons, like texts, can be read in various ways. *Some* readings we exclude; we think them wildly implausible and those who propose them become subjects for attributions of other, extrinsic further purposes. Yet among the plausible readings there are alternatives which, taken individually, are plausible and, even, insightful but which, collectively, cannot be accepted. All our resources of insight into the beliefs, attitudes, and circumstances and perspectives of others are called upon in such cases. Here we accept 'interim' attributions embedded in a wider view of what the agent 'must' have been doing in his circumstances *as we picture them*. Certainly some people are better than others at insightful and imaginative understanding of this kind;

some just do have more resources, just as some have greater artistic, mechanical, mathematical aptitudes. Psychological aptitude and competence just is not equally distributed among us. There are some whose views we, in actual practice, trust more than others; their accounts of the doings of others are deeper, more articulate, imaginative, more insightful and better organized, and so on. In describing these matters it is too easy to be banal and superficial. Wittgenstein's reminder from another context is relevant. In connection with describing our own practices, we are not prepared to describe *how* we act, though we act effectively. Similarly, we have not been trained in the ways of describing the greater psychological aptitude of some among us. But this does not argue against the reality of their greater competency. It is as much a matter of experience as anything could be. The objective psychological accounts and attributions of purposes, like the best accounts of historical episodes, works of artistic excellence, and other matters of importance to us, are not always obvious. Objectivity of attributions requires the reverse of dogmatism about the *particular* doings, character, and circumstances of others.

IV

Two accounts of the order of further intentional actions implicitly recognize the means-ends connections and the role of practical knowledge stressed above. Anscombe's account of the order of intentional acts (*Intention*, sections 23ff.) indicates that an action which is a fulfilled further intention (to poison the inhabitants, for example) of a previous or primary action (pumping) is the previously described action with the further intention as its purpose. Alvin Goldman's theory (*A Theory of Human Action*, 1970) is that where the means-end connection is operative, i.e., where an agent does one thing in order to do another, the end-act is "level-generated" by the former, means-act (assuming the further action or end-act is intended). (pp. 38–39) (Goldman does not wish to define level-generation, in any of its forms, in terms of means-ends connections since he is not interested in the generation of intentional acts at this stage of his theory. Here he is concerned with acts *per se*. I think it could be shown that his positive account of intentional action, with its qualification that the agent brings about an end-act *in a characteristic way* by another act, does implicitly involve the notion of means-ends connections, since the latter express the characteristic ways in which one act generates another. (Cf. Goldman, pp. 56–63) Both theories concern

themselves with fulfilled further intentions. Given that an agent did one thing as a means to another and given that the latter thing is brought about by the former in the way the agent anticipated or intended (it did not occur accidently, or *via* a wayward causal route for instance) both theories recognize that the fulfilled further intention is an act or action of the agent's.

But we have been considering successful purposive acts as well as things done with further intentions which were not fulfilled. A theory patterned on the former cases alone will not serve as an account of actions with further unfulfilled intentions. In spite of a failure or breakdown in the realization of a further intention, it is true that an agent does what he does with such further intention(s). In doing what he did (intentionally) he *was* also 'doing' (unsuccessfully) what further he intended. An agent may *be filling* the water supply tank in pumping although the tank is not being filled. A man may *be running* his auto engine in order to recharge the battery when the battery is not being recharged.

We can adapt the account of the order of further intentional doings provided by Anscombe. In doing A with the further intention to B, C, D, etc., one is performing D. (Cf. *Intention*, para. 26) But where the series is broken in that an intention beyond the intentional act A is not fulfilled, say D is not realized, still the reason the agent is doing (did) A is to do what further he intends, *viz.*, D. We may have a series of realized intentions and further intentions-with-which (A-B-C-D) when D answers the question 'Why?' about A and what the agent is doing (did) is D. This is Anscombe's sort of case. There may also be a series of partially realized further intentions-with-which (A-B-C-unrealized D) where D answers the question 'Why?' concerning A. In one case, in pumping, (A), he is filling the tank, (B), and thereby replenishing the water supply, (C), and poisoning the inhabitants, (D). In the second case he is pumping, filling the tank, and replenishing the water supply because of his further intention to poison the inhabitants, but he is not actually poisoning the inhabitants (for whatever reason, say, they learned of the plot and took measures to use another water supply). What he is *doing* intentionally is replenishing the water supply; what he is *intending to do* in doing this is to poison the inhabitants. What he is indending to do, poison the inhabitants, tells us why he is doing something intentionally, replenishing the supply, etc. His prospective action, (D), is as much a reason for his doing what he is doing as is his successful further intentional action. Here is the contrast: "Why did you replenish the water supply?" "To poison them" and he did. And, "Why did you replenish the water supply?" "To poison

them" but he did not, through no fault of his own or because some necessary condition was not present, etc. In the one case his reason is what he was doing and did; the other case shows his reason to be what he was doing but did not do. What he was doing, his further intention, is the same in the two cases.

V

Future intentions or intentions for the future should be considered a new concept relative to further intentions. Further intentions are intentions in acting; hence they presuppose the intentional actions which are the typical means to their fulfilllment. Future intentions need not be intentions in acting. One can always intend to φ only when Q (in the future) whatever one's current actions; one can intend to φ if and when Q (in the future); and one can intend to φ unless Q (in the future). One type of further intention, however, is only verbally distinct from a future intention. If I intend to A in order to *then* B, e.g., go early so I can *then* leave early, the intention to B is a future intention as well as a further intention in A-ing. While A-ing is not a means to B-ing, I believe I cannot B unless I A, that is, because I intend to leave early (B) and because of beliefs I hold about what is permitted of a guest, I believe I cannot, in one sense of 'cannot,' B unless I A. So *to B* is a further thing I intend in A-ing. In what follows I shall restrict the discussion to *intending to*, when this is understood to refer to a future intention as such.

Another respect in which the concept of future intentions is a new concept relative to the concept of further intentions comes out if we reflect that a child cannot come to have the concept of a future intention without having at least begun to master the future tense. A further intention *in* a present intentional action does not rely on a grasp of what the future holds, at least in primitively standard cases.

In order to have future intentions a child must have learned to perform or come to perform various voluntary actions and primary intentional actions. For intentional actions and refrainings are primary among the things an agent can intend for the future. This means that the verbs of action and descriptions of actions generally, since they are the linguistic versions of objects of intentions, are required for the expression of future intention. Thus, the concept of intending to . . . presupposes the concept of intentional action. Much of what we do intentionally we do without having *had* future intentions to do it. Hence, it would be a mistake to claim that intentional actions are realized future intentions.

Having intentions to . . . is not the same as expressing intentions to . . . Nor is formulating an intention for the future the same as having an intention for the future. To formulate an intention in speech or thought is an action; to intend to do something in the future is not an action, although it can be the effect of an action, when one intends to swim after throwing oneself into the river. "There might be a verb which meant: to act according to an intention; and neither would this word mean the same as our 'intend.' Yet another might mean: to brood over an intention; or to turn it over and over in one's head." (*Zettel*, para. 45)

Expressions of intentions for the future, in forms such as "I shall φ", can be fulfilled in standard ways by subsequent action. Failures to fulfill an intention exhibit considerable variety. Among these are: (1) The agent was lying and does not φ; (2) The prospective action of φ-ing is given up and the agent does not φ; (3) He φ's but not as he intended; (4) He undertook to φ but some one or more of the supporting conditions necessary for successfully φ-ing is not met. (5) He was forced to ψ, incompatibly with his φ-ing. (6) Supposing he truthfully said "I shan't φ" but he intentionally does what, in fact, is an instance of φ-ing because he does not foresee the way in which he will φ. Consider these in turn.

(1) If one lies in saying "I shall φ" he misrepresents his state or disposition. One who says "I intend . . ." may or may not intend as he says. What he avers is not false because, as it happens, he does not fulfill the content of the intention he avows. In the cases under consideration, falsity arises from lying. It may happen that one expresses the intention to φ, does intentionally φ, but was lying in avowing an intention to φ. This shows that it would be a mistake to identify expressions of intention with self-commands or self-exhortations. These last seem not to depend on states or dispositions of agents called commanding states. Commands and exhortations have performative features which expressions of intention lack. To be in a position to issue a command for J to φ, to say to J "I command you to φ" or, simply, "φ!" is to have issued a command, although it may not be meant or intended that J should φ. Nevertheless, to issue the command (intending its fulfillment or not) is to have commanded something. If this is true of self-commands as well, then it indicates a difference between these and expressions of intention, for merely to say "I intend . . ." or "I shall . . ." is not to intend something, even when it is said sincerely. For instance, there are illusions of future (and further) intentions, often in the form of self-deception. One may believe and say that one intends when one does not. Any theory must accommodate types of conative confusion and the failures in action they engender.

(2) An expression of intention for the future can misfire also because before it becomes mobilized at the time for action the agent alters his intention, abandons it, or adopts a contrary intention. Among the occasions for these changes are reconsiderations of the desirability of the further purpose of one's action, more drastic changes of heart so that what previously seemed appropriate now seems unacceptable to the agent, and recognitions of changed circumstances which for various reasons can render our further purposes unnecessary, ill-advised, too costly, etc. Not all alterations of intentions for the future are instances of changing one's mind. Suppose the future prospect was the purchase of some property featured as inexpensive, with a mountain view, at an altitude of 8400 feet, *viz.*, a certain parcel D. As it happens, the parcel of land designated 'D' is not inexpensive, has no view, and is at 6000 feet. D is no longer the object of interest when it is learned which parcel D is. However, E fits the description of the original intention and is, thus, a proper object of interest. "I shall purchase D, *viz.*, the one with such-and-such characteristics" is now amended by substituting 'E' for 'D' in the expression of the agent's intention. Here it can be said that the agent has altered his intention without, in one sense, changing his mind. Intentions like propositions one considers, can become clearer, become more specific, suffer some amendments without being abandoned.

(3) I truthfully say I intend to φ, and I have in mind a quite particular way in which this is to be done. I undertake to φ, do φ, but not in the way intended. I did what I did intentionally although I did not intend that it should be done in the manner in which it was accomplished. Fulfilling an intention for the future has among its conditions both accomplishment and manner of accomplishment. If the last is not met, the action is less than fully intentional. A less than fully intentional action is an intentional action. Consider: Jones intends to kill an antelope by shooting it in the heart with a single shot. Jones is permitted only one shot. He takes aim at a standing, nearby beast. He fires but the antelope moves, is struck in the head and is killed. Jones intentionally killed the antelope. His plan was not fully realized, however. Elsewhere we must consider the extent to which these phenomena are due to the essential generality of future contingent statements or propositions, the genus of which future intentions are one species. (See Chapter VIII)

(4) The view that fulfilling one's purposes is 'out of one's hands' once one has undertaken to perform a certain bodily movement actually expresses one of the variety of types of failures to execute an intention in acting. One fails to execute a plan when one or more conditions neces-

sary for its execution fails to obtain; the world or the gods do not cooperate with all of the best laid plans. Conditions outside the agent are not the only such causes of failing to fulfill an intention for the future or a further intention. Agent ineptness, disabilities, loss of skill, slips, and other types of familiar failures to execute are obvious.

(5) If an agent avows "I shall φ" and neither abandons nor amends his intention, does not change his mind, he may be forced to ψ, coerced in one or another way to ψ, where ψ-ing and φ-ing are such that they are not jointly satisfiable. The agent has kept to his intention; his failure may well consist in intentionally doing something incompatible with his initial intent, under duress, force, threat, or other forms of coercion. Here is a case of non-voluntary 'change of mind,' not a case of change of mind.

(6) Suppose one says truthfully "I shan't φ" and intentionally ψ's, which, in the circumstances, satisfied φ-ing. In that case the agent's avowal is truthful, he has not changed his mind, coercion is not present, but his belief that nothing he would undertake to do would be a case of φ-ing was false. He did not foresee the way in which he would φ. Promises can be 'broken' in this way; they are not exactly broken.

I have claimed that occurrent desires are neither necessary nor sufficient for having future intentions. Here is further evidence for this claim. (Even if there is a use of 'I want to . . .' which in certain contexts is equivalent to 'I intend to . . .,' still intending is not desiring). An agent may have always intended to keep secret what he knows about a friend, but it would be inappropriate to suggest that he continuously has desires or the desire to keep the secret. Or, another case, a person may intend to sell his inheritance to pay his debts, though at no time does he have a desire to sell what he inherited and neither does he have a desire to pay his debts. He must pay his debts or else; his only means is to use his inheritance, which he may do reluctantly, but not without intending to.

An agent may be almost irresistibly compelled by his desire to speak to a colleague about a matter but consistently intend to refrain from speaking of it. Or, another case, a person may fight his desire to insult another, a desire that occurs each time he encounters the other, but he may never intend to insult the other. His occurrent desire to X is not sufficient for his intending to X.

Nor is wishing to do something in the future or hoping to do it in the future the same as intending to do it. In general wishes need have no tendency to become realized. Some wishes are unrealizable, because what one wishes is impossible. Intentions for the future tend to be realized, necessarily. And to the extent that one is confident that something

is impossible to do, to that extent one will not intend to do it, as a matter of *psychological* fact. But a person may wonder if something is impossible and assure himself that it is (or that it is not) because he intends to try to do it and when the opportunity arrives, try to do it and fail. Having found it impossible to do, it will no longer be reasonable to undertake to do it, although he may intend to learn to do it or otherwise make himself capable of it. Even if *that* is impossible, he may wish it were not and hope that, somehow, it will become possible he may intend to do it if it is possible. Conditional intentions for the future are familiar. Conditional intentions remain distinct from hopings. One may hope to do something if it is possible without intending to do it if it is possible. For if one merely hopes to do it, if it is possible, he is not committed to setting himself to do it when he comes to think it is possible. If he fully intends to do it if it is possible and he comes to believe that it is possible, then his deeds fail to match his mind if he has an opportunity to try to do it but does not try to do it. No such failure follows from his merely hoping to do it.

The belief (or even knowledge) that something is impossible may be bracketed, not permitted its psychological and conative role, in cases of the following sort. A person is sure that it is impossible to move an enormous boulder. Others demur, suggesting that together they can move it. The disbeliever may enter into the attempt, trying as best he can together with others to move the object. His action is intentional. Their failure confirms his belief that the attempt must fail. So one needn't think something possible in order to attempt it. Attempting but failing is an excellent way to show that it cannot be done.

But intuitions seem to differ concerning the connection between intending to do something and knowing that it is impossible. As a matter of psychological regularity, one will not typically intend to do what he firmly believes he cannot do, unless his aim is to further support his firm belief. Apart from this last case, if an agent knows that his X-ing is physically impossible, does it make sense to say that he intends to X? Anscombe claims that "a man hanging by his fingers from a precipice may be as certain as possible that he must let go and fall, and yet determined not to let go . . . a man could be as certain as possible that he will break down under torture, and yet determined not to break down." (*Intention*, p. 93; Cf. G. H. von Wright, *Explanation and Understanding*, pp. 100–101; R. M. Chisholm, 'The Structure of Intention', *Journal of Philosophy*, LXVII, 1970, pp. 645f.) Both cases cited by Anscombe are instances of the internal negations of intentions. Intending not to X isn't equivalent to not intending to X, for in the former an agent would intend to refrain,

while in the latter he might have no intention on the matter. If an agent intends to X, then he cannot *not* intend to X, though it is not yet clear that he cannot intend not to X (For instance, he might not realize that X is identical to Y, while he intends to X and intends not to Y.) Clearly one can intend to X and intend not to X on different occasions. But can one both intend to X and intend not to X, in the same respect, on the same occasion? Let us call an intention conjoined with its internal negation a conjointly contrary intention. Is such an intention possible? A conjointly contrary intention is not logically excluded from the language; it is rather that an agent makes no sense to us if we say or we believe that he has a conjointly contrary intention. Such states seem possible; what we cannot readily do (or perhaps finally do) is reconstruct the intelligible considerations or practical reasoning of such an agent, if there was any, since we cannot see how carrying out a conjointly contrary intention could be a means to some intelligible end. The agent is unintelligible; we think he ought to be unintelligible to himself. But this does not logically exclude such a case.

The logical possiblitiy of intending to do what one knows one cannot do need not involve a conjointly contrary intention. A simpler case is just that an agent knows that he cannot withstand the torture and yet he intends to withstand the torture. Suppose he knows that he will try to withstand the torture, but also knows that in the end he must give in. This seems possible. Then does his knowing that he will try to withstand imply that he intends to withstand the torture? At the very least it seems to imply that he intends to withstand as long as he can.

CHAPTER FIVE

EXPRESSIONS OF INTENTIONS

> There might be a verb which meant: to formulate an intention in words or other signs, out loud or in one's thoughts. This word would not mean the same as our "intend."
>
> Wittgenstein

I

A characterization of the distinctive features of expressions and confessions of intention is now in order. Two kinds of cases are prominent. First, there are those in which an agent expresses, avows or accepts a further intention in acting. Anscombe's example remains useful: In operating a pump a man may be replenishing a water supply, and while it is true that the supply of water is being replenished, he is up to poisoning the inhabitants whose water it is, since he is pumping water with the further intention of poisoning the inhabitants. Or, another case of the same kind, in speaking to another person, one may be offending him, or although he was offended, the speaker only had in mind giving him some information. The man poisoning the inhabitants or giving some information may express, confess, avow, acknowledge or accept the attribution of such an intention; his language, if he uses any, will typically have a past reference, 'I did it (was doing it) to poison them,' or 'I said it to inform him.' Expressions of intention are used here to explain what was done, to put the agent's truthful view of it in the right light. Truthfulness here depends on the fact that the agent intended to do what he now expresses as the intention with which he acted.

The second kind of case consists of what Wittgenstein once dubbed 'momentary intentions' (*Philosophical Investigations*, 638), by which he seems to have meant mainly unheeded intentions which are expressed in such forms of words as 'For a moment I meant to insult him.' These, too, seem to make reference to the past. I shall be interested in investigating the question whether developments beyond the past episode to which reference in these is made and present developments, including

thoughts, beliefs, decisions, and reactions of the agent might not also play a role in the truthfulness of the expressions of such intentions. It seems clear that the history, before a past episode, is relevant to the attribution of this species of intention; I wish to support this point while also examining the question whether what happens, various details, reactions to and attitudes about one's past after the episode and before the expression or attribution might not also be relevant.

II

There are numerous points of contrast between the general conditions for the truth of descriptions and the truthfulness of expressions, avowals, or confessions of intentions. In this connection, Wittgenstein once remarked as follows:

> There is such a thing as colour-blindness and there are ways of establishing it. There is in general complete agreement in the judgments of colours made by those who have been diagnosed normal. This characterizes the concept of a judgment of colour. There is in general no such agreement over the question whether an expression of feeling is genuine or not (*PI*, p. 227).

This remark applies as well to expressions of intention, so that we may say that in general there is not *complete* agreement over the question whether an expression of intention is genuine or not.

Wittgenstein also remarks that:

> The criteria for the truth of the *confession* that I thought such-and-such are not the criteria for a true *description* of a process. And the importance of the true confession does not reside in its being a correct and certain report of a process. It resides rather in the special consequences which can be drawn from a confession whose truth is guaranteed by the special criteria of *truthfulness* (*PI*, p. 222).

This remark applies as well to truthful expressions of intention, so that the claim is that the importance of the true expression of an intention does not reside in its being an accurate report; rather, it resides in the special consequences which can be drawn from such an expression, confession or avowal whose truth is guaranteed by the special criteria of truthfulness.

In general there is complete agreement in our judgments of color. Indeed, in general there is complete agreement in a very wide range of judgments of identity, location, quality, quantity and numerous others. We do not disagree in calling *that thing* a cat, *that* a mat and, when they are situated like *that*, the cat is on the mat. So if we wish to speak truly we

should say that the cat is on the mat. The point, is, of course, that this goes for an indefinitely large class of cases beyond color judgments and things like cats on mats. We could give a drawn out story about this class and it would remind us that 'it is what human beings *say* that is true or false; and they agree in the *language* they use. That is not agreement in opinions but in form of life' (*PI*, p. 241).

Agreement, however, is not always complete agreement, nor is it always general agreement. That is, cases exist in which there is sometimes only partial agreement that something is so, and there are kinds of cases (some of them the same cases as before) in which only some who speak the same language agree. This is especially so in connection with topics that fall under the rubric of the philosophy of mind. While we can be as certain of someone else's sensations as of any fact (cf. *PI*, p. 224) there are conditions that would make room for doubt. As we move from another's sensations to his feelings, thoughts and intentions, agreement and certainty may persist although *room* for doubt may grow. Perhaps in attributing actions and intentions or expressing intentions and confessing to actions we just do go beyond what the narrow 'evidence' permits; the more complex, subtle, complicated, the action or intention, the more difficult it is to reach general agreement among those who share the 'evidence.' It is a feature of our mental life that judgments on which we agree run beyond anything we would call 'evidence.' Naturally, then, truthfulness and trust (a form of agreement itself) have important roles here.

Consider, for instance, one way of contrasting true reports of events and truthful avowals of intentions. Anscombe's example once again: a man would be speaking truly in saying that he is moving his arm back and forth and in so doing is operating the pump if he is in fact operating the pump. A man who said, in those circumstances, 'I am not operating the pump' would not be readily understood. That is not to say that he could not be understood, however. Perhaps he means to say that he is exercising his weakened muscles and it is irrelevant that the pump is pumping; he would do what he is doing even if the pump were not operating, and indeed, perhaps he would be moving his arm in connection with some other mechanism or even no mechanism at all, if the pump were not handy. And this may, of course, be true. Consequences follow from the latter that do not follow from his avowal that he is pumping water, and *vice versa*. If he is exercising, then we know that if we provide him with an exercising machine more convenient than the pump, he will not be uninterested from the point of view of practical considerations. Of course, he just might not be interested; then we would look for some explanation of

his lack of interest in our convenient exercising machine. If he is both exercising and uninterested in our machine, there is need for some explanation of his lack of interest that does not contradict his avowal that he is exercising. Further, if we make it clear to him that the pump is defective, he will not be interested if he is exercising rather than pumping; if we convince him that he will not get money for what he is doing, he will not mind, since he is not also pumping for pay in addition to exercising for strength. Each consequence we draw from an avowal of intention is, of course, open to debate, evaluation, further analysis, questioning and the like. Although we cannot specify it *a priori* and in advance it is typical that there is a point at which we will accept a person's account of what he or she is up to and on that there will be agreement. In this kind of case, it is as if we ought to say, at some point there will be agreement, unless there is not; but in the case of reports of what is happening or has happened, there will (categorically) be agreement. We rely on these types of agreements in *our* practices.

Thus a man may report that the cat is on the mat and it is there; there will be agreement. If a man says that he is pumping water to replenish the water supply and not just exercising, and he is moving his arm in such-and-such a way, there may be agreement or there may not. There are consequences of what he says, and we may proceed to examine them. At some point, there can be agreement. 'He is pumping, since he isn't interested in the exercising machine.' But, then, is he feigning lack of interest? Can he be caught 'off-guard?' What has he to say for himself? Is there genuineness in his voice? Is he clearheaded? Is this the sort of case in which he may suffer from or indulge in self-deception, sentimentality, scrupulosity and the like? If we doubt him, does our doubt arise from considerations about him or only ones about ourselves? (Cf. *Zettel*, 561–563). Barring defeats from these sources and perhaps others, there is a point at which we agree in our judgment; that is, we accept expressions, confessions or avowals of intentions on these terms, since they are *so far* our best explanation of his action.

Part of the point here might be put by adopting a convention for expressing the apparent difference that stress makes in such a negative avowal of intention as 'I'm not operating the pump.' (Cf. F. Dretske "Contrastive Statements," *Phil. Review*, 1972). Underscore the stressed element to indicate the locus of stress and the resulting contrast. Thus:

(1) I'm not operating the *pump*.
(2) I'm not *operating* the pump.
(3) *I'm* not operating the pump.

These are responses to the question 'What are you doing?' or 'Why are you doing *that?*' in the context in which it is clear that you are moving your arm back and forth in connection with the pump. (1) implies that the agent, in doing what he is doing, is not operating the pump, though he is doing something further in acting, and what he is doing is operating something, though not the pump (assume (1) is not conjoined with (2)). What he is doing in moving, etc., is, say, operating a signalling device disguised as a pump or connected in some way with a pump and that device operates whether or not the pump actually works. Or, maybe it works only if the pump does not actually work. His avowal could be 'checked' by considering these other facts and by considering his reaction to the information that the other device is not working. (2) implies that the agent is not operating the pump, although the pump may well be working as a consequence of what he is doing, or it may be working all on its own. Perhaps what he is doing is generating chemical changes in the muscles of his arm that have been found to occur when his arm moves in just the way it does when he is out to operate the pump or another gadget like the pump in respect to associated arm movement. (3) implies that it is not the agent who is operating the pump; perhaps the pump operates on its own and his arm is just along for the ride, for his amusement. Here he has no interest in operating something other than the pump; nor has he any interest in bringing about some other results which are produced by moving the pump handle as if to pump the pump. (2) has consequences different from, say (1). If (2) is truthful, he will take an interest in such matters as the experimenter's analysis of the chemical extracted from his muscle tissue. If (1) is true, he would be reasonably uninterested in such tests, though he is interested in the operation of the other mechanism. If (3) is true, he would find the request to remove his arm from the pump handle a generally reasonable one. Of course, one may pretend an interest or lack of interest. There may only be fine shades of behavior and general trustworthiness to go on. Some people may be better than others in judging these matters; and the agent's words will demand a special role if questions should arise about what he is really up to. Up to a certain point our judgment is revisable. But at some point for many cases we can be satisfied, for we are satisfied in may cases that a man intends what he does. The *logical possibility* of a better explanation is not a ground for doubting our explanation.

III

Someone might object that all 'the special criteria of truthfuness' notwithstanding, it remains true that the avowal or confession of intention we accept is still a report of the man's intention in the sense of a report of something *in* him, an event, disposition, thought, etc., which is his intention. This could amount to two things: either the claim is that what the man says or could say, e.g., "I am poisoning those captialists" has the form of a report or the claim is that what the man says reports an inner act of intending. The former is a claim about the 'clothing of the language' which I should not want to deny; indeed, it is just this that makes everything look alike. (Cf. *PI*, p. 224) The second claim could itself amount to at least two different things: (1) a truthful expression of intention reports inner phenomena, such as thoughts, images, feelings, where these may be thought of as inner events; and (2) a truthful expression of intention reports an inner speaking to oneself, a 'Let it be so' or an 'I will that. . .).'

Against the claim that expressions of intentions are reports of such things as thinkings, images, or feelings, let us imagine that another *knows* such details of one's inner life. So he ought to be able to say what I meant to do if he knows what thoughts, images or feelings I had. Indeed, he ought to be able to say with as much certainty as I what I was going to do or intended in acting. For if what I say in expressing an intention with which I acted is a report of the thoughts or images or feelings I had, and he knows what these are, then he should be able to issue the 'report' as authoritatively as I can. With this goes the reminder that if I had a multitude of thoughts, images, and feelings when I was X-ing, and in X-ing I was Y-ing with the intention to Y, someone who knew only my thoughts, images and feelings and that I was X-ing, need not know I was intentionally Y-ing – any more than he could know to whom I am referring if he knows only the images, feelings and thoughts in my mind. If God had looked into my mind, he could not tell from that whether I intended to Y; he, too, must at some point rely on expressions and surroundings as well as on what I *do*. (Cf. *Zettel*, 558.)

Futhermore, if expressions of past intentions are reports of inner events (not specified as thoughts, feelings, or images) then to remember what one meant to do would be to remember an inner event. Presumably, however, of any remembered *event* it can be said that one wants to remember it more precisely, more clearly and in more detail. When, however, we try to focus more precisely on the alleged inner event of intending, we remember, not that inner event, but other things such as

what else we were thinking and doing. This phenomenological argument holds only against an identification of what expressions of intention report (if they were reports) and inner events. In these cases it has been observed that the key thing, the inner event, seems to disappear when attention is focused on it, say, in the attempt to remember it *quite precisely*. Think of speaking with the momentary intention of deceiving; now one remembers having momentarily meant to do something. When one fails to find an appropriate inner event under the focus of attention, introspection or memory, one is likely to want to reply that we fail to adjust the focus properly. Accordingly, we looked at the wrong thing or in the wrong place. It is as if the past action, past situation, past thoughts, feelings and images *and* past momentary intention (conceived as an inner event) form a culture. The situation and action can be taken in at a glance or by inspection, but the rest requires a fine adjustment of the microscope of memory. But when we adjust the microscope what comes into focus is not the required inner event of intending. Rather, we find a variety, culture, of phenomena: what else we thought, what our feelings were, what we wanted, *and* present reactions to these. Of this Wittgenstein says, "Well, that only shows that you have adjusted your microscope wrong. You were supposed to look at a particular section of the culture, and you are seeing a different one. There is something right about that. But suppose that (with a particular adjustment of the lenses) I did remember a *single* sensation; how have I the right to say that it is what I call the 'intention?' It might be that (for example) a particular tickle is accompained by every one of my intentions." (*PI*, 646).

Suppose, however, I no longer remember my words but remember my intention in that situation. "I no longer remember the words I used, but I remember my intention precisely; I meant my words to quiet him." (*PI*, 648) Here, surely one must remember something very much like what a wordless-thought must be like when it is conceived as an inner event, since this seems to be what is remembered. Would it be like this case: "I no longer remember which machine or what special machine I used, but I remember my aim exactly; I meant to pound the crystal into powder." What memory shows here is not, as it were, just the past thought; rather, one has a sketchy picture of a past episode. "Exactly what the machine was like and how it worked is no longer clear to me. But it is clear that it was for crushing crystals." Memory may do no more that suggest words to me, "it was for crushing crystals" and so far that will count for my having had that aim. Perhaps other words will also be suggested that further sketch in the details of the picture of the situation I remember.

But none of these need be regarded as remembering an inner phenomenon. If I am inclined to say that I read off my intention from a memory of an inner phenomenon, that need mean no more that that I am *now inclined to say* that I remember certain states of mind and read off the intention from them. But what I 'read off' will either fit in and full out the situation or not; whether it does will be determined by my response and reaction now. If it is no part of my present explanation, my present reaction, it will not be an acceptable reading.

Consider cases of interrupted unrehearsed speaking. Later if we ask a person what he was going to say, it often happens that he can continue the interrupted sentence. Here one is likely to want to say that the thought he was expressing in speaking was already complete. To complete its expression is to remember the completed thought one had. Similarly, to be interrupted in doing something and to remember what one was doing in acting in that way is to remember the having of that intention or the awareness that such and such is what one was up to, and to express it now is to report an intention based on the memory of this awareness, experience, thought or whatever which is the having of that intention.

What now is this having of an intention or this awareness that one is doing Y in doing X? The question might be put like this: To what extent is one aware of what one is doing while doing it? One answer would be that one is aware of what one is doing while doing it just in so far as one doesn't only later realize what one was doing ('Oh, I didn't realize I was doing that!') but does know (can say) later what one was up to. (Cf. *Zettel*, 190.) If this is so, then to express what was one's intention is to exercise this knowledge or ability to say what one was doing. And to express one's intention in the form 'I was then about to . . .' is to exercise the ability which one has when one does not only realize later what one was then about to do and yet can say what it was. If, however, an expression of intention reports an inner awareness, experience, thought, etc., then, once again, if someone else were to know these thoughts, etc., he should know what one was then going to do; if he know what inner states the agent enjoyed, he should 'not just only later know' what the agent was about to do. Evidently, however, on the basis of such information about another's inner life, an observer would not often have this ability; often, at least, another person does only later come to know what an agent was about to do, and it seems perfectly consistent with this that he could have coincidental 'knowledge' of the agent's inner life. (Cf. *Zettel*, 44–45). This result is consistent with the possibility that there should be characteristic inner experiences associated with intending to do something or

being about to do something. Compare this remark of Wittgenstein's:

> "I was going to say . . ." – You remember various details. But not even all of them together show your intention. It is as if a snapshot of a scene has been taken, but only a few scattered details of it were to be seen: here a hand, there a bit of face, or a hat – the rest is dark. And now it is as if we know quite certainly what the whole picture represented. As if I could read the darkness. (*PI*, 635)

None of these arguments imply that we are uninterested in what thoughts, feelings and perhaps images accompany a person's action. Suppose that I become ashamed of an intention to deceive someone. "At that moment I hated him and for a moment I meant to deceive him." If I dwell on this I will no doubt remember what I thought about him, felt towards him, how he affected me, the conversation we had, what he said and my reaction and so on. If now I am ashamed of my intention, can I still feel satisfaction with the remainder of what I remember? The whole episode, indeed its whole history, seems to be what calls out my shame and gives me cause for shame. This history includes my thoughts, feelings reactions as well as my moment of hate and my intent to deceive. But we need not suppose that the momentary intention to deceive was itself another thought, feeling or episode alongside the rest. (Cf. *PI*, 642–644).

Another claim to be considered is that an expression of intention is a report of an inner saying; if an expression of intention, e.g., "I am leaving the room because you told me to," "I am replenishing the water supply" is a report of an inner saying, then the having of an intention is an inner saying.

A saying is something one issues; it is a type of doing. Accordingly, the question can arise Why are you doing that? And here the view under discussion is either in danger of a regress or else it involves falling back on another expression or formulating of an intention. An inner saying, "I am only pumping, not poisoning" is the kind of thing of which it can be asked "Why?" For instance, "Was that said with the further intention of distracting oneself from one's troubles, or said just because one simply always does say that when doing his job, or . . . what?" An intention can no more *be* an *inner* saying than in can *be* a *saying*, for if it answers the question Why? it cannot be just another one of the sort of thing for which the question Why? arises. More importantly, were the question Why? answered by reference to an inner saying, the answer need not succeed in connecting that saying with the relevant action. It could as well connect with conducting a thought experiment or hiding one's troubles as with the action of pumping. (Cf. Anscombe, *Intention* (1957), para. 27.)

Other considerations seem to weigh against the present claim. Assuming that there are collective aims and intentions in collective actions, for instance, in governments, institutions, corporations, communes and the like, and assuming that these would be expressed in plural sentences such as 'Our intention was . . .' it is hard to see what or whose inward sayings these sentences could report.

Furthermore, it seems clear that words that occur to a person, whether he utters them or not, could not be his intention. For at least he should have to *mean* them and not just have them occur in his imagination. "Meaning them" is sufficiently close to "intending" to indicate that we are really going in circles when we try to go to inner saying.

We could adopt forms of sentences which would make it seem that sayings were always expressed in expressions for further intentions-with-which and momentary intentions. We could preface our own expressions of intentions (or wishes, hopes, etc.) with the words "I said to myself . . ." and we might express the fact that someone else had an intention in acting by "He as it were said to himself 'I will . . .'." This would not make the confession and ascription of intention any more like sayings than attaching-the preface "I believe . . ." to ascriptions of pain makes real reactions to others in pain like expressions of mere belief or opinion. These verbal changes are possible, but nothing interesting follows for our problem, for we should ask next exactly how these new phrases are to be employed in actual cases, since "I said to myself" will not always mean that I spoke to myself and "He as it were said to himself" presumably will not mean that he said something to himself. (Cf. *PI*, 657–658).

We may conclude this section by noting that while the forms of words we use in expressing and confessing intentions-with-which and momentary intentions are similar to the forms of words used to report and describe events, they need not be thought to be alike in other, deeper ways. "The grammar of the expression I was then going to say . . . is related to that of the expression 'I could then have gone on.' In the one case I remember an intention, in the other I remember having understood." (*PI*, 660)

IV

There is sense of doing something such that whatever a person can be seen or known to be doing is something she or he is doing. Such doings may be intentional actions. We see that he is operating the pump, hear that he is speaking, learn that he is travelling to France. We could as well

see a motor operating the pump, hear a parrot speaking and read that the Chess Club is travelling to France. A human being can also tell us about an intention; and if he does so, we know something that we do not necessarily know in knowing that he is operating a pump, speaking to Smith or travelling to France. It is easy to think of this something more as inner, since it is about him and something in addition to what we can straight-off see, hear, or find that he is doing. Think of the case of momentary intentions. I say "For a moment I meant to insult him," when I did not insult him. What is the role of this kind of remark? Why should I tell another person this as well as telling him what I did and continued doing, in the sense of doing that he could notice and observe right-off?

Confessions of intention often serve some form of explanatory role. If such utterances do not refer us to inner events then when they are explanatory, the explanations evidently are not typical causal ones, in one familiar philosophical sense of "cause."

Expressions of momentary intentions set what one did in a wider or clearer context by revealing something of the agent which goes beyond what he did. They are ways we have for throwing the right light on what went on, which would not be present save for what we reveal in speaking as we do. Not that they merely interpret past actions and episodes, any more than "I thought this was the right size bolt" merely interprets my having tried and failed to use the right size. One need not have said or 'thought to oneself' that this bolt is the right size, in order that the later sincere utterance 'I thought *this* one would work,' *be* an explanation of one's failure. Explanatory thoughts and intentions explain and do not merely interpret actions.

The explanatory role of expressions of momentary intentions and expressions of thoughts ("I thought the bolt was right") may be brought out also by an analogy with explanations of meaning in the sense of reference. If I say "The man in the front row – I mean *that man*" (pointing) I am both referring and explaining my reference. In this sort of case, pointing is part of the technique of referring. (Cf. *PI*, 669) Contrast this with my more complex statement "When I said 'The man in the front row . . . I meant *that man*' (pointing). Here, too, what I say explains what I previously said because, except for my now lying, there is no more room for misfiring (factual mistakes, errors, mistaken recollections) than there is in the simpler, present tense case in which, while *pointing*, I say sincerely "I mean *him*."

The point of the analogy between specifying reference and expressing intention is that in the appropriate cases (1) specifying reference (by

pointing) gives the meaning of one's words, i.e., how they are to be taken by others and expressing intentions-with-which gives the meaning of one's action, i.e., how it is to be understood by others; (2) specifying reference later may give the meaning of one's previous utterance and confessing an intention later, such as a momentary intention, may put the previous action in its true light, so that others will understand what concretely happened.

In saying "For a moment I meant to insult Smith" (in particular circumstances) one may be engaged in explaining. Explaining here is consistent with the fact that confessing such a momentary intention can be part of a contemporary response to the past situation; the speaker now gives the right account of the past situation by way of a response to it. He has only later come to make a connection within a wider context, including such items as his previous dislike for Smith, Smith's harsh treatment of others, his attitude towards Smith in speaking to Smith, and so on. Hence, the explanatory intention appears in thought after the situation to which it is relevant. Wittgenstein remarks on this as follows: "Why do I want to tell him about an intention too, as well as telling him what I did? – Not because the intention was also something which was going on at that time. But because I want to tell him something about myself, which goes beyond what happened at that time. I reveal to him something of myself when I tell him what I intended to do. – Not, however, on grounds of self-observation, but by a reaction (one could also call it an intuition)" (*PI*, 659).

V

The description of the explanatory, clarifying role of expressions of past intentions has the air of paradox. Such expressions are about the past, yet they do not refer to the past, inner events, or, at least, not inner events of past intending, or, at least, not in every case. They explain one's past actions and attitudes but contain some important references to matters subsequent to the past episodes they clarify. How can this be non-paradoxically described and explained? Once again I shall attempt to make headway with the help of some remarks of Wittgenstein's.

'Grief' describes a pattern which recurs, with different variations, in the weave of our life. If a man's bodily expression of sorrow and of joy alternated, say with the ticking of a clock, here we should not have the characteristic formation of the pattern of sorrow or of the pattern of joy. (*PI*, p. 174)

'For a second he felt violent pain' – Why does it sound queer to say: 'For a second he felt deep grief'? Only because it so seldom happens? (*PI*, p. 174)

If I say 'I meant *him*' very likely a picture comes to my mind, perhaps of how I looked at him, etc.; but the picture is only like an illustration of a story. From it alone it would mostly be impossible to conclude anything at all; only when one knows the story does one know the significance of the picture. (*PI*, 663)

One thesis that seems to emerge from these and similar remarks of Wittgenstein's is that some emotions and intentions are not discrete events or episodes in the lives of people who have them. Rather, they refer to temporal patterns and historical scenes of experience and behavior without which they would not be what they are. The scene and surrounding story give these psychological and action relevant conations their significance, their meaning. They are nothing in themselves, apart from the weave and pattern of which they form a part. (Nor is a sound a phoneme apart from the structure and surroundings it has in a language.) Both the 'space' and the 'time' of many psychological and conative conditions contribute to the identity of these conditions. Their 'space-time' is the weave of our lives in which they have a role. This means, among other things, that whatever point in time, if any, such an emotion or intending has as its outset, what happens subsequently is involved in the individuation and identity of the emotion or intending. The same discrete moment which is the 'onset' of the emotion or time of the intention could be duplicated in two distinct sequences of action and reaction, differing patterns of action and reaction, such that only one sequence would contain the pattern or scene characteristic of the emotion or the subsequent developments characteristic of the intention in question. If such a view yields some illumination and removes the air of paradox it is due, in part, to the reminder that language and thinking function in a multitude of ways, some of which can seem paradoxical when compared with mere descriptive language and thinking. But we come close to substituting yet another air of paradox and one which points to deeper issues concerning the mental, for now it can seem that citations of intentions and emotions could not possibly have a clarifying and explanatory role, depending, it seems, for their significance on the conditions, events, and patterns of action they are said to clarify and explain. This deeper issue is treated immediately below and elsewhere; now the etiological and phenomenological side of the puzzle, later the theoretical side. (See Chapter Eight)

Expressions and confessions of intention are a vocabulary for our reactions and responses to what we and others do, have done, might do, think we will do, and purpose to do. They are truthful if defeating conditions

are avoided, e.g., if we are clear, sincere, genuine, not self-deceived, not overly scrupulous, and if the consequences they have for our beliefs, interests, other avowed and private aims, in short, the remainder of one's psychological condition, meet the tests that may be made. Such tests are, of course, dependent on our rough and ready generalizations and 'conceptual' knowledge of how psychological conditions hold together in the weave and pattern of life in more or less normal circumstances. We do achieve agreement as to the truthfulness of such expressions; we achieve a degree of agreement which excludes many possible explanations of what an agent was doing in doing what he did intentionally. In such circumstances, expressions of intentions are explanatory.

There is no short way of describing the possibility of the sort of truthfulness and common agreement which underlies the revelatory force of expressions of intention. What has been provided are only descriptions. Much more could be described in more or less detail: we doubt whether a person was doing what he brought about through his action, and we engage in speculations about what other things he could have been doing, given his circumstances, what we know or think him to be like, his beliefs, attitudes, and so on. The narratives of such speculations, carefully delineated, with analogies, etc., often just are acceptable. And some of these plausible accounts of what the agent was up to (in case he does not reveal it) will typically be accepted as what he could have been doing, given a rough plausibility ranking in light of what we know of the agent and the agent's situation. The most plausible narrative will be our account of what the agent was doing. Some other cases: another person disowns what this action conventionally means; we are unclear what we were doing in acting as we did and only later become clear by *now* formulating, verbally or in imagination, the plan of our previous actions; we come to wonder whether we intended or how seriously we intended all that we did in doing something intentionally; we formulate a plan in order to become clear about what we will do as well as to inform others what we will do. If one gives a detailed description of these phenomena, phenomena where the roles of intending and intention expression vary and contrast in their variety with simpler descriptions and reports, then the explanatory and clarifying force of expressions becomes obvious at the level of particularity at which agents think, possess attitudes, live, and act.

Suppose we ask how it is that human beings ever came to make the verbal utterances which are here called expressions of wishes and intentions. Indeed, we should be struck by the existence of such a practice as,

for example, confessing to past intentions! Why is that in the whole diversity of mankind we do not find many who don't exhibit this practice? Perhaps only the feeble-minded never come to use such expressions as "I was on the point of . . ." or "I was just going to . . ." (Cf. *Zettel,* 43). How should we imagine humans learning these uses of words? They were not taught these in just the way they acquired words for colors and familiar objects. There may be a clue in the analogy between speaking and doing and the analogy between meaning and intending. Children first learn to speak rudimentary parts of the language and later to say what they mean or what they meant. Others react to them and eventually they respond, sponaneously, with such utterances as "I didn't mean *that*; I meant *this*" and they tell us. Similarly, children act and also we teach them standard actions and activities and later to say what they are doing and what else they were doing in acting. Then they spontaneously respond to the actions and reactions of others with confessions of intentions of their own. Mostly people mean what they *say*; mostly people intend what they *do*. But they also react with "I didn't mean *that* (what I said)" and "I didn't intend *that* (what I did);" both are based on reactions to what happened, i.e., to what they and others took them to be saying and what they and others took them to be doing. So human beings come to have verbal expressions for these reactions because they want to *tell* us something about themselves which goes beyond what happened at the time, (Cf. *PI*, 659). We evidently have a *need* for self-explanation. The citations which satisfy part of this requirement include expression of what we've called "momentary intentions."

Our conclusion is not that there are no grounds for saying "For a moment I meant to deceive him." The conclusion so far is only that no momentary inner event is my evidence for "For a moment I meant to deceive him" and that the surroundings in which this is said are not my evidence for saying it. Of course something makes it possible for me to say it, namely, that for a moment I meant to deceive him. Such grammatical facts are not in question. A good way of thinking about them is at issue. Thinking of my "meaning for a moment to deceive someone" as expressing a momentary mental event whose occurence makes it possible for me to say now "For a moment I meant to deceive him" combines what is a perfectly correct but grammatical point, *viz.* that if I did not intend to deceive him, then I'm not entitled to say now that I did, with a wrong picture of the use of "For a moment I meant to deceive him," *viz.* that this describes a momentary, inner event. Rather, its role, exhibited in human responses and language, is to explain what one *did, didn't do*, and one's moral attitudes and relations.

CHAPTER SIX

INTENTIONS: THE STRUCTURES

> It is at least worth considering, in what different senses, we are said to do what we did, and what we did not design to do . . . Sometimes we mean by action the determinations or volitions themselves of a being of which the intention is an essential part.
>
> Richard Price

I

'Intentionally' is an adverb. It can be treated as a sentence modifier, verb modifier, or prepositional phrase. Thus, "John jumped" can go into "Intentionally John jumped," "John jumped intentionallly," and John jumped with the intention to jump." For simple cases there is no reason to insist on one or another reading. For complex cases a more complex structure is evidently required. Thus, "John jumped onto the bed to escape a spider" would, on the simple readings go into

(1) Intentionally John jumped onto the bed to escape a spider.
(2) John jumped onto the bed to escape a spider intentionally.
(3) John jumped onto the bed to escape a spider with the intention to jump on the bed to escape a spider.

But each of these is ambiguous. One possible reading of (1) does not imply that John's intentions was to escape a spider. The speaker of (1) identifies what it is John jumped to escape, *viz.* a spider, though John only saw something he intended to avoid but did not recognize that it was a spider. Further, suppose John jumped, intentionally, onto the nearby covered piece of furniture, not knowing that there was a bed under the covering. John did not intend to jump on the bed (rather than anything else) intentionally, though he intentionally jumped onto what was, in fact, a bed. Reading (2) has one clear interpretation on which John did what he did intending to escape a spider. It was a spider John intended to escape by jumping onto the bed; or, intending to escape a spider, John jumped onto the bed. Paraphrases such as these highlight or stress the various

elements modified by forms of 'intend'. Sentence (3) contains all the ambiguities by simply repeating the action sentence in the prepositional object phrase. No doubt pragmatic features help make clear what the target sentence means.

Emphasis, stress, pragmatic presupposition can each add to disambiguating these sentences. We can, however, express the differences with some 'structural' additions. By so doing, we can indicate the possible readings without recourse to pragmatic and conversational considerations. Implications and presuppositions can be explained with the aid of the notion of range or scope employed in uncovering these structures. Other data are also represented in this way.

An essential feature of intentionality is that actions, further reasons and intentions for acting, and objects acted on, are intentional only 'under certain descriptions.' The intentionality of the objects of intentions and of action can be displayed with the notion of the range of intentions and the structures it permits. Since the intentionality of the objects of intentions is crucial to the assessment of intention expression and subsequent action, this notion of structure has a useful job to perform. Again, related data and puzzles (e.g., the intending 'paradox') can be accommodated in this way. Finally, the structural features of intention negation will be examined below.

II

Consider these groups of sentences.

(A1) Oedipus intentionally struck the rude old man.

(A2) Oedipus' intention was to strike the one who was, in fact, the rude old man.

(A3) Oedipus' intention was to strike the one he saw (*viz.* the one in front of his eyes) who happened to be the rude old man on the road.

(B1) Oedipus intentionally struck the person (who was a rude old man) with the intention of injuring him.

(B2) In intentionally striking the one who was the rude old man, Oedipus intended to injure that person.

(B3) Oedipus' intention was to injure by striking the man who happened to be the rude old man on the road.

(C1) Oedipus intentionally struck, with the intention to injure, the old man who was, as Oedipus recognized, rude.

(C2) Oedipus intended that the rude person should be injured by Oedipus' intentional striking of him.

(C3) Oedipus' intention was that the rude person should be injured by Oedipus' striking because Oedipus recognized that he was rude.

Contrast the groups (A), (B) and (C) with

(D1) Oedipus intended that the rude old person should be injured by Oedipus' intentional striking of him (and, of course, the one he struck was a man).

There are various differences among the groups (A), (B), and (C) and the entry (D1). Each exhibits some uses of some forms of 'intention.' Importantly, the range, scope or diffusion of forms of 'intention' increases or expands from the entries in (A) through sets (B) and (C) to (D1). This *range* difference can be displayed. These differences are logically and pragmatically and psychologically significant.

Let us adopt the following convention: Use *iota* subscript to indicate that what immediately precedes it is in the range of (or is modified by) an appropriate form of 'intention.' By abbreviating and applying this convention to (A) through (D1), we will write 'Ov $_{\iota}$' for 'O intentionally v-ed'. 'Ov $_{\iota}$ (φ) $_{\iota}$' will display 'O intentionally v-ed with the further intention of φ-ing'. For sentences of the form 'O intentionally v-ed with the further intention of φ-ing someone or some thing he took to be or recognized as an S' we write 'Ov $_{\iota}$ (φ) $_{\iota}$ S $_{\iota}$'. (As we note below, if O does what he does because, in part, he recognizes or thinks of S in the manner he does, then we will need another, clearer, way to display this role of 'intends'.)

Thus a representative from each of (A) through (D1) becomes

(A1) Oedipus struck $_{\iota}$ the rude old man.
(B1) Oedipus struck $_{\iota}$ (to injure) $_{\iota}$ the rude old man.
(C1) Oedipus struck $_{\iota}$ (to injure) $_{\iota}$ the rude $_{\iota}$ old man.
(D1) Oedipus struck $_{\iota}$ (to injure) $_{\iota}$ the rude $_{\iota}$ old $_{\iota}$ man.

Sentence (A1) is not true if Oedipus did not know or suppose that what he was doing was striking. While a mindless brute, say, can strike something or even strike out at something, it cannot strike $_{\iota}$ anything, since it is absent the concept or notion of striking. What Oedipus can do $_{\iota}$ is limited by what he can know, think or consider. For example, because he was not in a position to know that he was moving molecules about in striking the

rude fellow, Oedipus could not have been moving $_{\iota}$ molecules in the vacinity. Oedipus did move a few molecules in the neighborhood when he struck $_{\iota}$ the rude old man. But he neither considered this fact, counted on it, or knew that he was doing such a thing as this.

(A) is true, by contrast with (B1) even if Oedipus does not know that the one he struck $_{\iota}$ is a man and is old and rude. Normally, of course, if (A) is true, so is (B). Normally, one will have a reason for striking intentionally, and one's reason may well have to do with features one attributes to or recognizes in the thing one strikes $_{\iota}$. Still, (A1) and (B1) differ in an important respect as noted.

Oedipus' striking $_{\iota}$ of the rude old man may have been done with a further intention or purpose in mind. If Oedipus struck $_{\iota}$ (to injure) $_{\iota}$ the old fellow, then Oedipus acted with a further intention, since we take it that Oedipus knew or supposed that striking is a way to hurt the one struck. Here, then, is a sentence of the group (C): "Oedipus intends to strike, with the further purpose of injuring, the one he is striking." Schematically: 'O s $_{\iota}$ (I) $_{\iota}$ x' where x is the rude old man.

If Oedipus intended that the rude (old man) should be injured by Oedipus' striking of him for his rudeness, Oedipus' action was fully intentional. The sentence which is used to express what Oedipus' did is a fully intentional action sentence. These are sentences of the form of (D1). Schematically, they are 'O s $_{\iota}$ a $_{\iota}$,' where 'a' names or designates the rude old man.

So far we have noted the variable range of forms of 'intention' over verbs of action, adverbial modifiers and phrases, and complements and objects of verbs. What of the agent-subject of intention sentences? As far as so indicated, instances of any of the groups (A)–(D1) could be true though Oedipus did not know that his name was "Oedipus." So one might suggest that intentional action sentences of each sort should contain something like 'Oedipus $_{\iota}$ . . . etc.' This topic is of major importance. It is fully covered elsewhere in our discussion (See Ch. Seven pp. 109–110). For present purposes let us assume that our specimen sentences could be reformulated as, say, 'Oedipus knowingly says: "I intend to strike . . . etc.",' reflecting the presence of the indirect reflexive. Thus, we read 'Oedipus v $_{\iota}$ (φ) $_{\iota}$ a $_{\iota}$' as Oedipus intends that he himself v . . . etc.'

To recapitulate, fully intentional action sentences assert that it is someone or something *qua* an identifying or characterizing description to which something is intentionally done with a further intention of doing that. For instance, 'Oedipus intends (intended) to injure the old man *qua*

rude in striking him intentionally.' Accordingly, Oedipus must be credited with the conception of rudeness if such a sentence captures his full intentionality. Below, this feature will be represented as the notion, concept, name or description of the object acted upon by displaying the object *under* that notion, description, etc. (Oestension will be treated separately.) So, 'd'/a represents the object, a, as recognized or thought of by the agent. If 'd' is iota-subscripted, then the agent intends his action on a-as-featured-by-d. If 'd' is not iota-subscripted, the agent only recognizes *a* as, thinks of it as, is aware that it is, d or a d, though such thinking has no role in the agent's reasoning, planning, or intention formation. Here we must distinguish three grades of intentional objectivity. The first is represented by Ov_{ι} a, the second by Ov_{ι} (d/a), and the third, full object-intentionality, is represented as Ov_{ι} $(d/a)_{\iota}$. In English, these are (i) "Oscar intentionally took a step (and stepped on the line)." (ii) "Oscar intentionally took a step and (he noticed) stepped on the line." (iii) "Oscar intentionally took a step intending to step on the line (and did so)."

This account of the diffusion or range of intentionality will be employed to replace less fine grained accounts.

III

'Oedipus struck the rude old man who was his father' appears to entail 'Oedipus struck his father'. But 'Oedipus struck $_{\iota}$ the rude old man (the one who was Oedipus' father)' appears not to entail 'Oedipus struck $_{\iota}$ his father'. The account of the diffusion of intention accounts for these data.

The truth conditions for (C1), 'Oedipus struck $_{\iota}$ to injure $_{\iota}$ the rude $_{\iota}$ old man', include Oedipus' knowing or believing that what he undertook was a case of striking and Oedipus' undertaking was aimed at someone he thought of as rude with the purpose of injuring that person. If we deny that (C1), thus understood, entails 'Oedipus struck $_{\iota}$ his father', our denial does not yet precisely indicate what we wish to exclude. For consider the following sentence pairs, (where the rude old man is Oedipus' father)

(Ai) Oedipus struck $_{\iota}$ to injure $_{\iota}$ the rude $_{\iota}$ old man.
(Aii) Oedipus struck $_{\iota}$ to injure $_{\iota}$ his father.

(Bi) Oedipus struck $_{\iota}$ to injure $_{\iota}$ the rude old man.
(Bii) Oedipus struck $_{\iota}$ to injure $_{\iota}$ his father.

(Ci) Oedipus struck $_{\iota}$ to injure $_{\iota}$ the rude $_{\iota}$ old man.
(Cii) Oedipus struck $_{\iota}$ to injure $_{\iota}$ his father $_{\iota}$.

Entailment certainly fails in the (C)-pair. Since the verb of action, the modifier of purpose, and the object are all within the range of iota-subscripts in (Ci), (Ci) reports a *fully* intentional action. Suppose (Ci) is true. Fully intentional action presupposes the performance or occurrence of its less fully intentional elements. Accordingly, (Ai), together with the sameness of the rude old man and Oedipus' father, entails (Aii). But in the (C)-pair, the truth of (Ci) does not presuppose that Oedipus' father, the rude old man, must be so conceived by Oedipus, as a condition for the truth of (Ci). Hence, the (C)-pair differs from the (A)-pair in just this respect. Entailment fails. (Cii) is certainly consistent with (Ci). But, again, Oedipus' recognition of the one he struck $_{\iota}$ as his father is not required for either his striking $_{\iota}$ the rude old man or his striking $_{\iota}$ his father, i.e. the one who is, in fact, his father. The (B)-pair is unproblematic.

Although the rude old man is the same person as Oedipus' father, the rude $_{\iota}$ old man is not in fact the same as Oedipus' father $_{\iota}$ in Oedipus' reasoning, thinking and acting. The recognized-rude-old man is not the-recognized-father in Oedipus' view, thinking, planning or acting.

Intentional activities normally have ordinary objects and persons as objects. The attribution of a fully intentional action is, typically, by means of a sentence or statement which, when true, requires that an object is either identified by the agent of the action or is thought of in some way by the agent. Only thereby is the agent's identification or thought conatively active. Indeed, full intentionality is not by dint merely of an identification, thought, or description of the object by the agent; further, such a thought also plays a role in the agent's psychology in relevant plans and acts (if any). This is the typical, normal situation of human, fully intentional actions.

In extraordinary situations intentional action is still possible. For instance, one might find an agent 'grasping an hallucinated dagger'. Such an intentional action is *necessarily fully intentional*. The act of grasping a dagger is contingently intentional in that one can both grasp a dagger, not intending to do *that*, and one can have intentionally grasped a dagger. By contrast, 'grasping an hallucinated dagger' is necessarily intentional, just in that there is no room for unintentionally (e.g., unknowingly) grasping an hallucinated dagger. In such extraordinary cases, the intentional object of the intentionally modified verb of action (to grasp, in this case) is exhausted by the psychological role played by the description, thought or identification on the agent's part. Necessarily, necessarily fully intentional actions of this sort are extraordinary. Unless it was typical of agents to do such things as intend to and succeed in grasping such things

as daggers, no one could be credited with the occasional 'grasping an hallucinated dagger.' Thus, in the usual case fully intentional actions will have intentional objects of action to which actual objects correspond. Hence, from $Ov_{\iota}(\varphi)_{\iota} a_{\iota}$ it follows that there is an object (actual object, person, etc.) and it stands in a relation to a_{ι}, here called 'corresponding to'.

Less abstractly, if Jones correctly identifies what he is doing something to, *corresponding* will be a case of instancing. That is, the description under which Jones acts on *a, viz.* d/a, will hold of *a*. But instancing is by no means sufficient for the full intentionality of Jones' deeds perpetrated upon *a*. As well, the description or thought by virtue of which *a* corresponds to the intended object of Jones' action must play its role in Jones' conative states and processes. (The issues involved here are more fully discussed below, in Ch. Nine).

IV

The present account of intention modifiers easily explains a variety of other data, among which are principles of the diffusiveness and heredity of intentions noted by H-N. Castañeda. (See his 'Intentions and the Structure of Intending', *The Journal of Philosophy*, LXVIII, 1971). Castañeda formulates the following Diffusiveness and Heredity Principles.

> *IN.2 'X believes that his (intention) "I will A" implies his (intention) "I will B", and X intends to A' *implies* 'X intends to B'.
>
> *IN.3 'X intends to A and X believes that the fulfillment of his (intention) "I will A" requires causally that his (intention) "I will B" be fulfilled' *implies* 'X intends to B'.

These alleged implications hinge on how we analyse the intentions 'his (intention) "I will A"', 'his (intention) "I will B"', and the propositions 'X intends to A'. In *IN.2 and 3 the objects of the intentions as well as the verbs of intention and the further intention are not explicitly represented. Since these elements of intentionality were relevant to the question of entailment discussed above, we may suppose they are relevant to implications claimed in *IN.2 and 3.

Consider *IN.3. Using the present notations, three versions of this Principle are available. Modify for the future tense with '(future)'.

(I) Oedipus strikes $_{\iota}$ (future) the rude old man, *and* Oedipus be-

lieves that his striking $_{\iota}$ (the rude old man) causally requires that his (intention) to disturb $_{\iota}$ molecules in his neighborhood be fulfilled *implies* Oedipus disturbs $_{\iota}$ molecules in his neighborhood.

(II) Oedipus strikes $_{\iota}$ (to injure) the rude $_{\iota}$ old man, *and* O. believes that his striking $_{\iota}$ (to injure) the rude $_{\iota}$ old man causally requires that his (intention) to disturb $_{\iota}$ molecules $_{\iota}$ in his neighborhood be fulfilled *implies* O. disturbs $_{\iota}$ molecules $_{\iota}$ in his neighborhood.

(III) O. strikes $_{\iota}$ (to injure) the rude $_{\iota}$ old man *and* O. believes that his striking $_{\iota}$ (to injure) the rude $_{\iota}$ old man causally requires that his (intention) to disturb $_{\iota}$ molecules $_{\iota}$ in his neighborhood be fulfilled, *implies* O. disturbs $_{\iota}$ molecules in his neighborhood.

(I) and (II) claim that there is an implication from an intention and a belief to an intention. While in (I) there is a claim that the intention in the antecedent of this implication is a partially intentional action, in (II) the intention in the antecedent is a fully intentional future action. The consequent in (I) is a partially intentional future action; in (II) the consequent is a fully intentional action for the future. And in (III) we find an alleged implication from a fully intentional future action and a causal belief to a partially intentional future action.

(III) is pretty clearly true, since a fully intentional action statement (as in the antecedent of (III) presupposes its corresponding less fully intentional actions (as in the consequent of (III)). On both (III) and (I) the subject of action, Oedipus, is credited with an understanding of what he intends to do and is credited with a further intention. Thus, he intends to strike to injure, and he intends to disturb (molecules) in order to move some of them, etc. Naturally, therefore, Oedipus intends to do these things, i.e. the verbs of action are intentionally modified. So (III) and (I) are obvious.

However, (II) appears to parallel the pair (Ci) and (Cii) in which implication fails. But his is not a genuine similarity. (II), unlike (Ci) and (Cii) is a true implication, since (II) credits the agent, Oedipus, with a causal belief according to which his future striking $_{\iota}$ the rude $_{\iota}$ old man causally requires his disturbing $_{\iota}$ neighboring molecules. No doubt this is a most peculiar belief; yet it is a possible one. If Oedipus believes thusly and, more centrally, if Oedipus relies on the truth of his belief in his pro-

ceedings, i.e. in his planning and acting, then (II), unlike the (Ci) and (Cii) pair, is an implication. And this is because the (Ci) and (Cii) pair are silent on matters of the agent's further beliefs and the roles of those beliefs in his practical thinking.

*IN.2 and 3 hold on the present account of the diffusion and range of intentions, in their variety of modifying roles. Since both of these are plausible principles and since the discussion of them illustrates the way in which the present view accomodates intentions-wth-which, further-intentions, and now intentions for the future, our results are added evidence in support ofthe usefulness of the distinctions and differences expressed in the notion of the intentions' range.

We can generate schemata of implication based on the ideas so far discussed. For example:

(1) A v_{ι} (to φ) $_{\iota}$ a_{ι}
implies
(2) A v_{ι} (to φ) $_{\iota}$ a
and (a=b) implies
(3) A v_{ι} (to φ) $_{\iota}$ b
implies
(4) A v-ed (and φ/ed) a
implies
(5) A v-ed a
implies
(6) A v-ed

In such a schema the principle of implication is that iota-subscripts may be dropped from right to left. And non-iota-subscripted terms thus introduced can be replaced by substituting an identically designating term.

Further, some consistencies among intention-modified action sentences yield logically weaker principles. These are principles of intention introduction. For instance, if an agent intentionally v's for the further purpose of φ-ing something, then there is *typically* some description of the thing the agent v's, and the thing in question is the intentional object of the agent's v-ing. Three principles arise from such a consideration.

(T1) If the left-most term of an action sentence falls within the scope of the iota-subscript, then its right-hand neighbor tends to do so as well. Example: Typically, if A v_{ι} a, then A v_{ι} φ_{ι} a. English: Intentional doings typically have accompanying further intentions or purposes.

(T2) If the left-most term and its right-hand neighbor both fall within the range of the iota-subscript, then if the object term is the right-hand neighbor of a term in the range of iota-subscript, then the object term tends to fall within the range of the iota-subscript. Example: Typically, if A $v_\iota \varphi_\iota a$, then A $v_\iota \varphi_\iota a_\iota$. English: Typically, an agent knows that to which he is intentionally doing what he is doing intentionally, and he intends to do *that* to *it*.

(T3) If A v-ed, typically A v-ed a, for a wide variety of v's and a's. English: Typically, we intend what we *do*, where 'intend' is understood to fall somewhere or other in the continuum of intentional activities.

Any number of writers have noted further data (the so-called accordian effect) which can be explained in terms of the notion of the diffusion of intention. (See Joel Feinberg, 'Action and Responsibility' in M. Balck (ed.), *Philosophy in America*, 1965; Donald Davidson 'Agency' in R. Binkley *et. al.* (eds.) *Agent, Action*, and *Reason*, 1971.) Davidson describes a case which illustrates this phenomenon.

> A man moves his finger, let us say intentionally, thus flicking the switch, causing a light to come on, the room to be illuminated, and a prowler to be alerted. This statement has the following entailments: the man flicked the switch, turned on the light, illuminated the room, and alerted the prowler. Some of these things he did intentionally, some not; beyond the finger movement, intention is irrelevant to the inferences, and even there it is required only in the sense that the movement must be intentional under some description. In brief, once he has done one thing (move a finger), each consequence presents us with a deed; an agent causes what his action causes. ('Agency', p. 16).

It is clear that this sort of case can be fully expressed in the iota-subscript notation. For instance, if the agent, Smith, intentionally moves his finger, intentionally flicking the switch thereby, then Smith moves $_\iota$ his finger $_\iota$ to flick $_\iota$ the switch $_\iota$ and this event causes Smith's illuminating of the room. But accordians must both expand and contract. So we can add to this effect the typicality phenomenon already noted. For instance, if the agent, Smith, intentionally moves his finger, intentionally flicking the switch thereby, then Smith intentionally illuminates the room (the friendly world agreeing). Again, plausible intuitions are perspicuously displayed in the iota-subscription notation.

V

I have insisted throughout that intention formation, intention adoption and intentional action presuppose action competency, in the form of sheer physical ability as well as acquired practical capacity with the ways and means of things. Such powers are expressed in voluntary actions and primary intentional actions which are the natural and learned means for doing things characteristic of most normal agents. Agents acquire capacities and abilities in social life, which is a culture with its ways of doing things and its things to do. Natural powers are assigned significance in such a culture. It is not possible, consequently, that an agent such as Socrates could have intended, on a certain occasion, to read *Hamlet* or to put out the milk bottles. In the normal capacities and among the voluntary actions of Socrates nothing would be these things to do and nothing would be means for doing them. What can be done, what is done, what is to be done, what can be intended, and what further can be intended are partially a function of historical and so cultural circumstances of agents, and the actual world in which they live, plan, and act.

Even the actual objects of simple future intentions will have a limited number of possible intentional descriptions or descriptions under which they are objects of future intentions. An actual object cannot be brought under a description which contains concepts foreign to the agent who has an intention whose object is that actual object. Alexander might have had in mind to ride a certain horse in Persia; he could not be said to have in mind to ride a certain quarter horse or to have in mind to move some molecules in his neighborhood, although the last are always things *to* which future intentions are directed when their actual objects are made of them.

Further expressions of future intentions present two aspects: predictive futurity and intentional futurity. Thus, "I shall φ" predicts (i) φ, (ii) φ brought about by me, and (iii) φ brought about by me in a certain manner. Depending on the details of the case, an expression of intention cannot stand without revision or criticism if not-φ at the appropriate time. Nor can it if φ, but φ was in no way brought about by me (at the appropriate time). And if φ was brought about by me in some unanticipated or some unheard of way at the appropriate time, what happened does not exactly fit what I said I intended to do. So our target expressions have at least these three parts to which a natural order of inquiry corresponds: Is it that φ? It is that φ because of the agent? Is it that φ because of the agent in the appropriate manner?

The non-predictive aspect of expressions of intention (the aspects are distinguishable but not distinct in psychological reality) resides in the fact that they express the decisions, desires, choices, values, tendencies, expectations, in short, the *mind*, of the agent. Expressions and manifestations of future intentions are in no sense performatives; they do not produce what they are about. The agent lies if his words do not reflect his mind, his thinking what to do, and the like. Insincerity and lies are predominate ways in which expressions of future intentions can misfire. Mistakes can, however, also occur. A lie is present if the agent avowing 'I shall φ' either does not have it in mind to φ or is still in doubt whether he will undertake to φ. He implies that he has no doubt that he will undertake too, so he is not a truthteller if he is in doubt, is undecided, whether he will undertake to φ. The possibility and nature of mistakes, in contrast to insincerity, will be discussed elsewhere. Suffice it to note here that it is not obvious how a sincere mistake of the appropriate kind would be diagnosed. Mental and emotional confusion, not knowing one's own mind in one or more of the variety of ways this can befall one, are among the possible explanations. What I wish to examine here are cases of mistakes in the way the objects of intentions are featured, since our notion of $_{\iota}$-modification and its scope can help illuminate this phenomenon.

VI

What happens when a reference to or description of an object of a future intention contains an error on the part of the agent expressing the intention? For example, suppose an agent avers "I shall paint that window next" when in fact what he is indicating is not a window. There are several possibilities. He intends to paint *that* expanse or area next; whether or not it is a window is not actually relevant. He will, he is thinking, paint *that* next, and he is indicating it as a window. Or, he intends to paint that *window* next; he is engaged just now in window painting, say. When he discovers his mistake in thinking that *that* is a window, (it is only a picture of a window), he will not paint it, of course. So his intention can be directed on and to that, whether or not it is a window, or it can be directed to *that* but only on it as a window. In the last case, the description of it as a window will be essential to it as an intentional object of action. In the first case, "it is a window" is merely the agent's way of indicating an area or object which is the object of his intention. Using the notation previously introduced I will indicate this difference as follows: $JV(P)_{\iota}(d_{\iota}(w)/a)$ where $'d_{\iota}(w)'/a)$ is read as 'the object, area, expanse *a,* under the refer-

ence *that* and which is a window.' Contrast, $JV(P)_{\iota}\, d(w)_{\iota}/a$ where '$d(w)_{\iota}$'/*a)* is to be read 'that area, object, or expanse, *viz.* the window.' In the case of $JV(P)_{\iota}(d_{\iota}(w)/a)$, it is false that *a* is a window, but "*a* is a window" falls outside the range of the intention of the object of intention. Hence, the falsity does not alter the appropriateness or identity of the object for the agent's purpose, plan or future intention itself. But if the agent actually *V*-ed *a*, having had the intention expressed by '$JV(P)_{\iota}\, d(w)_{\iota}/a$ and '$d(w)_{\iota}$' does not hold of *a* since *a* is not a window, then he did not do what he intended *to what* he intended to do it to. He did what he intended, that is, he did some painting, but he did not paint what he intended to paint, that is, the window. On the other hand, if the agent *V*-ed *a*, having had the intention expressed by '$JV(P)_{\iota}d(w)/a$ then he both did what he intended and did it to what he intended, although he incorrectly indicated what that was. So, for suitable values and with suitable assumptions,

(I) If $JV_{\iota}(d(P)_{\iota}/a)$ and $\sim (Pa)$ and *JV*-ed *a*, then JV_{ι}-ed $(d_{\iota}(P)/a)$ and $\sim JV_{\iota}$-ed $d(P)_{\iota}/a$. If Jones intended to paint the window and he painted what he designated as the window, then, though Jones intentionally painted what he designated, he did not intentionally paint a window.

(II) If $JV_{\iota}(d_{\iota}(P)/a)$ and $\sim (Pa)$ and *JV*-ed *a*, then JV_{ι}-ed $d_{\iota}(P)/a$. If Jones intended to paint what he took to be a window, *viz.*, that area and he painted that area, though it was not a window, then Jones intentionally painted what he intended to paint.

The duality possible in the objects of simple future intentions – the duality between the essentiality of the identifying description and the inessentiality of it, also affects the further intentions, i.e., the *in-order-to* features of further intentions. For the further intentions that are present in future intentions are to be fulfilled in the manner or way anticipated or expected by the agent. Their fulfillment *simpliciter* will leave it open whether the expression of the future intention, in its predictive aspect, is accurate. For the expression of a future intention which contains a further intention (e.g., "I will φa in order to ψ) four basic conditions of success are possible, with one basic condition dividing further. In typical cases, the agent can be said to (I) perform the action(s) which fulfill the future intention (the means) and succeed in the further intention (the end) in the manner anticipated. Or: (II) While the agent fails to act as he intended or as his expression predicts, somehow the further intention expressed in his avowal of a complex future intention, his purpose, does get realized. Case (II) will divide into cases of accidental success, cases of

only indirect success and cases of success by another means. In condition (III) and (IV), the predictive aspect will be unfulfilled, for the agent will have failed in his attempt (Case III) if he acts but his further intention is not achieved, and the agent will have failed altogether (Case IV) if he neither acts nor is it the case that the end is achieved.

Cases (II) and (III) require discussion, since case (I), simple success and case (IV), complete failure, are reasonably obvious. Condition (II) divides, first, into the sort of situation in which the agent does not do what he intended as means, but a sufficient condition for the materializing of his further purpose is satisfied nonetheless. His intention was to read the document in order to identify Mr. X; he did not read the document, and he failed to do anything else as a means to discovering who X is, but the identity of X was revealed to him in a conversation he overheard at a café or in a dream or was written in the clouds. His intention, when judged against what he did, why he did it, and against the upshot, was a failure, or perhaps, at best, an accidental success. (Among the cases of this sort there will be relevant moral considerations that support the withholding of any praise for such successes.)

Secondly, we can recognize the sort of situation in which the agent partially fails to do what he intended as means to fulfilling a further intention, although what he does do turns out to have been sufficient for his further intention. The agent intended to labor for a certain amount of time, since that amount was thought necessary for a certain wage and was sufficient. He failed to work the required time, but got his desired wages from an improbably beneficent employer. Here the agent's undertaking is only indirectly successful, since the end was only indirectly a consequence of the agent's actions.

Thirdly, cases are imaginable that resemble the first special case of Condition II, in that the agent does not do what he intended as means, nor does he do part of what he intended as means (unlike the last example) but he does another thing which turns out to be sufficient for his purpose. He intends to insult a certain person by saying a certain thing to him in a certain manner. Instead, he says nothing and thereby insults.

So we can formulate these situations as follows:

(I) If $J.V_\iota\varphi_\iota a_\iota$ and JV's, V-ing causes φa, and V-ing causes φa in the manner anticipated, then J's action, predicted in $J.V_\iota\varphi_\iota a_\iota$ is a success.

(II) (i) If $J.V_\iota\varphi_\iota a_\iota$ and $\sim JV$'s and φa (via another causal route), then J's undertaking, predicated in $J.V_\iota\varphi_\iota a_\iota$ is an accidental success.

(ii) If $J.V_{\iota}\varphi_{\iota}a$ and $\sim JV$'s but J partially V's and φa (via another causal route to which J's partial V-ing is relevant but not sufficient) then J's action, predicted in $J.V_{\iota}\varphi_{\iota}a_{\iota}$ is indirectly successful.

(iii) If $J.V_{\iota}\varphi_{\iota}a_{\iota}$ *and* $\sim JV$'s and JU's φa and J's U-ing is sufficient for φa then J's action, predicted in $J.V_{\iota}\varphi_{\iota}a_{\iota}$ is successful by another means, a happy upshot. (Assume V-ing = U-ing.)

(III) If $J.V_{\iota}\varphi_{\iota}a_{\iota}$ and JV's and $\sim\varphi a$ (V-ing was not sufficient for φa), then J's action, predicted in $J.V_{\iota}\varphi_{\iota}a_{\iota}$ is a failure in the attempt to φa, an unsuccessful attempt.

(IV) If $J.V_{\iota}\varphi_{\iota}a_{\iota}$ and $\sim JV$'s and $\sim\varphi a$, then J's action, predicted in $J.V_{\iota}\varphi_{\iota}a_{\iota}$ is a failure, a complete failure.

Another example: Jones intends to steal money from the store belonging to Smith in order to get funds to repay a debt to Brown. Suppose that, all unknown to Jones, Smith just sold the store to Green, who has deeded it to Jones himself, and Brown, perhaps impressed with Jones' new status, has forgiven Jones' debt, so that there is no store which belongs to Smith, the store Jones has it in mind to rob is his store, and Jones has no debt he owes to Brown. Evidently, since Jones can't steal from himself, he can't steal money from the store in question. And since no money he gets can be used to repay a debt he has been forgiven, the in-order-to connection has been broken and can't be realized. He can, of course, take money from the store in question with a *false* view to giving it to Brown in payment of the no longer existent debt. But what he does, in fulfillment of his intention, will not be stealing. Nor will what he does be either necessary or sufficient for removing his debt to Brown, since it no longer exists.

The present discussion has proceeded on the supposed idealization that future intentions were plans, purposes and designs on actual objects or on future states-of-affairs with ingredient actual things. This is not really much of an idealization, however. There is only a bit of truth in the idealization claim. It seems that agents can and do intend to do things to what is, in fact, nothing at all. For example, you now fully intend to drown your cat's *next* litter. Though you don't have the next litter now, at least you have the cat. but suppose you intend to drown your next cat's first litter. This requires that you believe that you'll have a cat, in fact, an unspayed female cat. For surely there is some sort of nonsense in "I fully intend to drown my next cat's first litter but I doubt that my next cat will have kittens." Perhaps "I've decided that if I have another female, I'll do in her litter, but I doubt I'll have another cat" is not peculiar. One can decide to do something if something else and doubt that one will do it

because one doubts that the condition will hold. But intentions for the future, strictly so-called, do not permit such a doubt. Where they seem to do so, intending merges with, or just is, hoping or wishing; it is something less than fully intending in that it can only be mobilized if and when its conditions are fulfilled (or are believed to be fulfilled). But our full intentions, one might say, go up against the actual world and propose what we take to be possible states of it. Of course, we can be mistaken about what is actual as well as about what we can do. But general skepticism is not possible for normal agents whose activities make a difference.

VII

These structures can show what is happening in the so-called intending analogue of the Good-Samaritan paradox. (Cf. H-N.Casteñeda, *Thinking and Doing*, 1975, pp. 167–168 and Ch. VII, Sect. 10). The problem arises in the psychologically peculiar case of Jones who knows that he is driven and compelled to hurt the feeling of those he encounters during the period of a full moon. He knows, too, that he will encounter his friend, Smith, on the evening of the next full moon (call this 10/10.) Worried about his contact with Smith on 10/10, Jones decides that on 10/11, the day after the full moon encounter with Smith, he will do something special for Smith when he (Jones) recognizes that he has hurt Smith's feelings on 10/10 *without intending to do so*. Jones' current intention can be expressed as:

(I) Jones intends to do something special for Smith on 10/11 when he recognizes that he has unintentionally hurt Smith's feelings on 10/10.

We notice that, (II), "Jones wil do something special for Smith on 10/11 when he recognizes that he has unintentionally hurt Smith's feelings on 10/10" *implies* "Jones will hurt Smith's feelings on 10/10." Suppose, (III), that Jones knows or believes (II). Now, it might be proposed, as a principle of 'intention implication' that (Pi): If X intends A and X believes that X's doing A implies X's doing B, then X intends B. (Pi) has the *look* of a plausibe principle. But from (Pi) and (ii) and (III) it follows, (IV), Jones intends to hurt Smith's feelings on 10/10; yet (I) implies that Jones does not intend to hurt Smith's feelings on 10/10.

How is this result to be avoided? Plausibly, we should reject (Pi.) This can be seen to be plausible when we express (I) in the suggested notation. Thus, (I) is;

(a) $J_\iota D_\iota (h/s)$

(a) is not the same as

(a') $J_\iota D_\iota (h_\iota/s)$

in that (a) does not mean that Jones intends to do something special for Smith, having intentionally hurt Smith's feelings; (a) does *not* contain the act (h) of hurting Smith's feelings within the scope of the i-operator. But since in (a) s occurs under the description h, that he has hurt (or will have hurt) Smith is something Jones believes; Jones thinks of s-ing in this way. (a) does not say that h is what Jones intends s as; but in (a') h is in the scope on the i-operator. And although (a') implies (a), the reverse is not true. (Pi), in effect, warrants the incorrect inference from (a) to (a'), since it, in effect, warrants the implication from (a) to (a') type propositions. Hence, (Pi) should be rejected.

In place of the rejected (Pi), we can propose the principle

(Pib) If X intends A and X believes (or knows) that X intends A implies X intends B, then X intends B.

Thus,

(b) $J_\iota D_\iota (h_\iota/s)$

and

(c) J believes that (b) implies $J_\iota D_\iota (i_\iota/s)$

then

(d) $J_\iota D_\iota (i_\iota/s)$

Thus, if Jones intends to do something special for Smith after he (Jones) has intentionally hurt Smith's feelings by insulting Smith (intentionally), then Jones intends to, among other things, insult Smith.

VIII

Standard propositional negation is normally explained as a function which takes propositions into propositions. In addition, however, there is the distinction between 'external' (normal truth functional) and 'internal' negation. The last is sometimes associated with 'predicate negation', so that we have (1) "a is P" becoming "it is not the case that a is P" by external negation, (2) "a is P" becoming "a is not P" by internal negation, and (3) "a is non-P" as the predicate negation of "a is P". No doubt

the full account of the differences and connections among (1), (2) and (3) will require an account of the sort of term "a" is, among other issues. These issues are not our present topic. Rather, we can approach the question of intention negation by looking at the question of the parallels, if any, between propositional negation and intention negation.

A typical view is that, whatever the ultimate fate of the distinction between external and internal propositional negation, intention negation does not permit the distinction. On this view, there is no distinction between "It is not the case that I intend to Q" and "I don't intend to Q". It appears, however, that the last can be true either because the speaker has no intention with respect to Q in that he has both no intention to Q and no intention to not-Q or because while he has no intention to Q, he does intend to not-Q. We might think that in the form "I shan't Q", the speaker is both denying any intention to Q and affirming an intention to not-Q, while "It is not the case that I intend to Q" carries not implication that the speaker intends to not-Q. However, it is certainly not clear that this is a ground for rejecting the typical view, since (i) it might well require a stipulation concerning "I shan't Q" and its implications and (ii) it is not yet clear what connections hold between either "It is not the case that I intend to Q" and "I don't intend to Q" on the one side, and "It is not the case that I intend to Q" and "I shan't Q" on the other.

Consider (ii) first. Are 'I intend to Q" and "I shall Q" (where 'shall' is not the 'will' of mere futurity, but the intentional future) equivalent? No doubt in context

(1) I intend to Q

and

(2) I shall Q

can be used with the same force and implications. Syntactically, of course, they certainly differ. "Intend" can take that-clause complements, whereas "shall" does not. (1) has such paraphrases as

(3) I intend the following: to Q

and

(4) I intend that I shall Q,

where (4) is not equivalent to

(5) I intend that I intend to Q.

Furthermore, (1) appears to be interpretable either as a relation or

notionally. Thus, the relational reading of (1) is

(6) There is something such that it is a case of Q-ing and I intend it.

Of course, the internal structure and components of "Q" will yield a more complex version of (6) when they are fully specified. If "Q" is, for instance, "take the dog for a walk", then (6) is, roughly (beyond first-order)

(6′) There is something, x, such that x is a dog and there is something, T, such that it is a taking for a walk, and I intend to Tx.

The notional reading of (1) is something like

(7) I intend that there is something such that it is a case of my Q-ing,

or

(7′) I intend that there is an x such that x is a dog and that there is something, T, such that it is a taking for a walk and I make x to satisfy T,

or some such.

This distinction suggests one condition on an account of intention negation, *viz*. if (1) is read relationally, as something like (6), then intention negation had better be read as internal negation. It is questionable that the external negation of (6) or (6′) makes sense, i.e., that both there is no such thing as Q-ing and I intend Q are possibly true. On the other hand, if (1) is read as (7) or (7′), the question whether intention negation is such that the internal/external negation distinction collapses simply reappears. Does the negation of (7) or (7′) yield 'I intend it is not the case that I Q' or 'I intend it is not the case I make x to satisfy T'? These are the questions with which we began the excursion into the relational/notional interpretations of (1), so no help has been gained by this excursion, with the possible exception that the relational reading is not very plausible.

Let us return directly to (ii). And let us take it that "It is not the case that I shall Q" is the same as "I shan't Q" where the last expresses the speaker's intention to not-Q. If this is correct, then since it is at least questionable whether "It is not the case that I intend to Q" is the same as "I intend to not-Q', we seem to have uncovered an important difference between "I intend to Q" and "I shall Q". Although we might continue to think of both as intention expressing, they will now be different sorts of intention expressions insofar as their negations differ, the first permitting, evidently, two conditions for negation, the second permitting only one. That is, "It is not the case that I intend to Q" is true in case either the speaker has no intention with respect to Q or no intention with respect to

not-Q, whereas "It is not the case that I shall Q" is true just in case the speaker intends not to Q, intends not-Q-ing, intends to not-Q. Thus, what the issue in (ii) suggests is that the typical view – intention negation collapses the external/internal negation distinction – holds for intentions properly expressed only by "I shall . . ." but does not hold for those intentions only properly expressed by "I intend to Q".

This result suggests that intentions themselves are of two types. Those expressed by "I shall . . ." and only thereby, we will call "mobilized intentions". These are roughly comparable to what have been called "volitions"; intentions to do some thing, here, now. They express the speaker's setting himself to do something, here, now; they are the outset of sophisticated agency. Thus it becomes clear why the external/internal negation distinction fails in such cases; setting oneself to do some thing here, now has as its contrary in the circumstances of action, setting oneself to do some *different* thing. Hence, contrasting decisions are naturally expressible by "I shall Q" and "I shan't Q" where "shan't Q" is satisfiable by "not-Q-ing" (if the contrary is itself intrinsically contrary to "Q-ing") and by "R-ing" (if in the circumstances of action "R-ing" precludes "Q-ing").

On the other side, intention expression (like expression of belief) can be *express*, i.e., explicit, as in "I intend to Q" (and "I believe that P"). The expressness of this form consists in part in its relative context independence. Whether a saying of something without the express "I believe that . . ." or the explicit "I intend . . ." is an expression of belief or intent will depend on features of the context and the speaker's purpose. Whether a saying of "I shall . . ." or "I will . . ." is an expression of intention (rather than of prediction, of a promise, of a threat, and so on) will depend on features of the context and the speaker's purpose. But "I believe . . ." expresses a belief, if candid. And "I intend . . ." expresses intention, if candid.

"I shall . . ." is particularly suited for the expression of so-called mobilized intentions or volitions. (Nothing at this point will hang on how much I will be stipulating in contrast to describing usage.) Volitions are indexed to the present time, place, and circumstances of the agent. Like present orders (e.g., "Fire!", "About face!" "Put it there!") they are creatures of the present. It is suited to them that they get expressed, as intentions, in unqualified, direct, simple, and free-standing forms. (Compare the furtherance of the speaker's purposes in selecting to use "Fire" rather than "I order you to fire.") In direct but non-express intention expression, with "I shall . . ." and "I will . . .", the infinitives (e.g., to φ) are not used. Noun forms for the intended actions are employed.

Thus, "I shall φ", not "I shall to φ". Exclusive predicate alternatives are appropriate, since the agent is either going to do φ or not do φ in the actual world of action. And mobilized intentions are not iffy, wheny, or in-casey. Intentions for the future, by contrast, are suited for expression in *express* forms, with infinitive constructions; here attached conditions are appropriate. (To stretch the point, consider "I intend to Q just in case P but not before R when S".) Thus, a brief study of intention negation suggests data we have noted elsewhere, *viz.* the differences between explicit, 'clothed' intention expression, and direct, non-express, 'naked' intention (volition) expression.

It seems that "He will Q" is reserved for such things as predictions, guesses, prognostications, and the like. "You will Q" (or "You shall Q") seems more or less reserved for similar purposes and for threats and the like. The denials of such expressions are, accordingly, either negations of such predictions or negative predictions (or expressions of threats not-to-do something). Trivially, neither affirmatives nor negatives of these third and second-person forms are *expressions* of intentions of the referents of the personal pronouns in them. One can describe and can attribute another's volition; one can urge another to (or demand that another) adopt an intention for present action; one can manipulate another into willing something here and now. But what one cannot do is adopt such an intention for another. Hence, no expression or utterance is an expression of another's intention. For the same reason, nothing counts as denying another's first-person intention (intention negation is essentially first-personal since intention is essentially first-personal). I can, of course, deny that you have an intention to do something; I can deny you the opportunity of fulfilling such an intention; I can attribute a negative present intention to you; and I can express my belief that you are unable to fulfill your present intention. What I cannot do is express or deny *your* mobilized intention, your intention is the ostensive mode.

Similar points hold for express intentions for the future and for standing intentions. I can attribute a disposition, plan, a purpose to you. In doing so, I will, of course, attribute first person reference to you. "John intends that he marry a wealthy and unliberated woman" can be my way of attributing the intention John would express by "I intend to marry . . . etc. . . ." If I deny that John has such an intention, I attribute to John something he would express by "I don't intend . . ." (if I am correct in my attribution).

The typical view that there is a collapse of the internal-external negation in expressions of intention is supported by these data.

CHAPTER SEVEN

INTENTIONS: THE CONTENTS

Causa finalis movet non secundum suum esse reale, sed secundum esse cognitum.

I

The aim of this chapter is to study some possible accounts of the nature and role of intentions (contents) in light of an array of data we must first describe. Any plausible view must accomodate these data. We will examine some leading candidate theories. The view which emerges in the end is eclectic and lacks the unity each of the theories examined attempts to feature. This is due to the force of the complex data. A simplistic theory can seem plausible only because some of the data are ignored, so that we get a truncated version of the role of the contents of practical thinking, considering what to do, and intending to do some thing. We must attempt to balance the complexities we encounter with the need for a simple theory of the nature, structure and role of (even) simple intentions.

II

The data are exhibited by describing the way intentions and contents of conative states (action relevant states) are conceived in common sense psychology. I will be describing, then, what I take our view of this aspect of ourselves and our thinking to be.

Ask a person what she intends to do, now that it is evident to everyone that what she has just been doing has not succeeded in getting the results she has undertaken to accomplish, and she may well say something like "Well, now I think I shall try this" and she may give a description of what she intends to do now. Intentions are given by descriptions of what the agent thinks will be means to set ends in such a case.

Ask a person what she intends to do, now that it is evident to everyone that some purpose she has attempted to accomplish is not attainable, and she may well say something like "Well, given that *that* can't be managed, I'll accept the next best, and do *this*" and she may give a description of what she now accepts as worth undertaking. Intentions are given by de-

scriptions of ends, goals, things thought desirable by the agent, in such a case.

Hence, expressions of intention include descriptions of means to be employed for further purposes or intentions as well as descriptions of further purposes or goals themselves. Both plans and purposes are expressed as forms of intention. Schematically, an agent may intend to do X with the purpose of doing Y; and an agent may be said to intend to do Y (by some means or other, including merely Y-ing if that is all that seems required). Two forms for such expressions are available, *viz.*, "I shall X in order to Y" and "I shall Y". In the last the intended means are not explicit. They may be obvious, given the agent's powers and the nature of Y, e.g., Y is putting the breakables on the top shelf. They may be problematic given the nature of Y, e.g., Y is getting John to study extra long for an examination. ("I shall somehow get John to study extra long for that exam.") They may also be problematic because of the limitations of the agent.

Under what conditions, if any, can such contents be said to be true or false? It has been remarked by various theorists that "I shall φ" when it expresses an intention for the future is not either true or false, unlike, for instance, "I will φ" or "He will φ" when these are mere predications of what the speaker or referent of 'he' will do. They are correct, accurate, true in whatever way predictions admit these predicates. But what of the "I shall φ"? (Nothing important turns on whether 'will' or 'shall' expresses mere futurity or intentional futurity; the distinction is evident enough for present purposes. I choose 'shall' for intentional futurity; the question is whether this is different, and if so, how is it different from the 'will' of mere predictions. Compare Wittgenstein's case: "I shall take two powders now and I will be ill in one hour.")

It seems clear that an expression of intentional futurity, whatever its other features, is also a kind of predication. This is based on the consideration of such conversational scenes as: "You said you intended to φ! What happened, you didn't do it? Did you change your mind?" "No, I didn't change my mind; I said I intended to to do it; I didn't say that I would do it!" This reply is not coherent, not even if we qualify the responses with "would try" rather than the pure success form "do it". One who candidly expresses an intention, e.g., is not reading the sentence "I intend to (or shall) φ" in a short story, etc., and who has not changed his mind (or intention or decision, etc.) is misusing language with the reply "I didn't say I would φ". (There may, of course, be an intention behind this misuse in the particular case.) One way to describe the misuse of

incoherence is to claim that an expression of intention is also a claim that the speaker will do the thing. And if he doesn't change his mind and neither does the thing nor undertakes to do the thing, then something is wrong with what he said in saying "I shall do it".

Some confusions on the question of the truth or falsity of expressions of intention arises from the view one takes on the truth or falsity of predictions themselves. One wants to ask, here, when is a prediction true or not true? If we take predictions in the sense of things thinkers issue, make, set forward in their public or private speech, and so, in other words, as if they were (at least) a kind of performance which makes a claim about a future contingency (let us say), *then* with these restrictions we are not in danger of denying excluded middle or bivalence if we say that predictions, like prognostications, prophecies, guesses and so on are correct when and in case the thing happens. We use 'true', 'not true' where we might as well without loss use 'correct', 'incorrect', 'came true', 'failed', 'is as you said', etc. Before the future develops one way or the other, we withhold verdicts on predictions and the like. Those of us with very strong realist views about truth may still have it that, so to speak, behind but not in the prediction or prognostication there is a proposition, and it is either true or false. But this strong realism *cannot* be a consideration for us, given that we take predictions in the way proposed here. And this way of taking them seems apt if we are concerned with similarities with expressions of intentions for the future. Behind this are strong realist intuitions without much confidence that 'universal determinism of the future' is any clear idea; a picture perhaps but not a clear idea.) *If* we take prediction in the way suggested (in rather the way Ryle once proposed, for instance) there will be less reason to find puzzlement over the question of the 'truth' of the predictive aspect of intention expressions. Less does not amount to zero, however.

I suggest then that expressions of intention ("I shall . . .") and sayings that one intends ("I intend to . . .") are correct in their predictive role, as sayings about what one will do oneself, in case the future is as one says, i.e., in case one who does not change his mind does undertake to do what he said he would. If undertaking it is all that is required of him to do the thing, then his expression is correct when the thing is done by him. We can as well use "true", "is as you said", "did as you said", etc., in place of "correct"; what is to be remembered is that these have the role "correct" has as a verdict on what happens when it has happened or else did not happen. This way of taking the matter caters to the datum that one person can issue a prediction concerning what an agent says he intends to do,

what it is he expresses an intention to do, since the predictor and agent seem to be 'talking about' the same thing, *viz.* what the agent shall do. This way of taking the matter does not exclude the possibility that there is an epistemologically inaccessible *proposition*, with the agent as its subject and some action as its predicate, and *it* is true or else it isn't. But, again, such a proposition cannot enter any of our considerations. Further, such proposition might be true of future events (the actions which satisfy their predicates). But if as suggested elsewhere in these studies (Cf. Chapter One, pp. 5f) freedom of indifference is a presupposition of the application of our intention notions, then the possibility of either fated events *or* universal determinism of action-events cannot be a consideration for us, even if either is realized.

The notion of *truth* which seems most central for intentions might best be called "truthfulness." In the expressions of intention there can be *truthfulness* or its absence. An agent is truthful in connection with his intentions if he has what he expresses; if he has decided to do, has adopted the intention to do, and has not altered his decision or rejected the intention to do X when he expresses it, then he is truthful in expressing his intention to X. His expression itself is truthful, i.e., matches his intent, if he is not lying and makes no verbal slips in its expression. There are complications here already noted (Cf. Chapter Five, pp. 81–83) Confusion, self-deception and other conditions which render a person and his conative states semi-opaque to him seem possible; he may not be in a position to issue a truthful expression of his desires, decisions, and intentions. Whatever else they are, intentions can be opaque to the agent to one degree or another.

Intentions themselves, in contrast to their expression, are evidently not either true or false. Although they will prove to have sentential–like structure, involve reference and predication, they are not sentences or porpositions, not either true or false. Yet they are consistent with, inconsistent with, imply and are implied by other intentions. They are conjoinable with others of the agent's intentions; there are disjunctive, biconditional, and conditional intentions. So these features, in the data, must be combined in a theory in which they are not, as such, bearers of truth. This tension in the data, as here described, must be accounted for in a plausible theory of the nature of intentions. No doubt we will see that the descriptions of data will themselves need to be refined as we begin to assess theories of the nature of intentions.

We have noticed elsewhere (Cf. Chapter Three, pp. 40–41) that intentions can be generic, containing action predicates (universals) and inten-

tions can be singular in a strong logical sense, e.g., intentions in the ostensive mode. Their predicates can, of course, have adverbial modifications beyond the intention to . . .when . . ., or the intention to π at time t, etc. The temporal references themselves can be general or specific, e.g., when I find he has done . . . or at exactly high noon on July 4, 1989, in Dry Creek, Wyoming.

Our theory of intentions must also make them suitable contents for the roles they have in the account of the dynamics of agency. They must express plans and purposes; they must be fit for fulfilling or realizing to one or another degree. They must suffer the possibility of whatever degree of 'match' with undertakings and actions is required for fully intentional action. We should cater to standing intentions in the case of 'life plans' and dispositions of agents throughout various spans of their lives and activities. And, of course, intentions must be fit as contents of conative states and dispositions; they must be 'objects' fit for entertaining, adopting, and fulfilling, and they must be the kind of content which is generated by other cognitive, conative and emotional conditions.

The relational or wide versus notional or narrow scope reading of first-person intention sentences does not represent a psychological, conative difference. Since intentions are essentially first-personal contents, this distinction makes no difference. Naturally it makes a difference in the attributions of intentions. Thus, "I intend to follow the spy when I see him on the beach" and "There is a spy on the beach and I intend to follow him when I see him" are not different intention contents. The following are different: "There is a spy on the beach and John thinks he sees him and intends to follow him" and "John intends to follow the spy he thinks he sees on the beach."

Logical exportation does seem permitted in connection with both expressions and attributions of intention, however. "I shall tell him . . . as soon as he arrives" seems to common understanding to imply "There is something I shall do as soon as he arrives". "He intends to V *a*" seems to imply "There is something he intends to do to *a*". Indeed, in normal contexts common understanding seems to have these yield "He is such that I shall tell him . . . as soon as he arrives" and "*a* is such that he intends to V it". Strong *assumptions* of actuality and existence are involved in the normal understanding of intentions and intending. Common sense also recognizes 'purely intentional' objects of intentions, e.g., intentions whose ingredient objects are hallucinated. (Cf. Chapter Six, pp. 89–90) By emphasis, tone, and other signals intentions directed on non-actual objects can be attributed. "Poor man, he intended to write down the names

of the ladies he *said* he *saw* in the garden." Here I am imagining the case of a victim of a cerebral stroke hallucinating women he thinks he recognizes in the garden and setting out to write down their names. Should the victim say "I shall write down the names of the women in the garden" and we see that he is hallucinating we can attribute and report his intention accurately only with a gloss of some sort indicating the hallucinatory status of the objects of his intention. Yet we do not deny him this intention. It is not that he thinks he intends something or other; rather, he thinks there is something or other concerning which he intends . . . .

There is also a non-personal notion of intention. Suppose one asks of part of an unfamiliar mechanism "what does this handle do?" In explanation it may be said "It was intended to operate as a brake." Here the idea of a subject does not appear. (Cf. Wittgenstein, *Zettel*, 48.) Further, subjects do not appear in intentions when we speculate about intentions, wonder what anyone who is the first to arrive on Mars will consider doing first, what his alternative intentions will be, and wonder what it would be like to be in a position to consider the intention to initiate a thermonuclear war, wonder what was the intention in some ancient contruction and so on. Here we are considering intentions without any idea of the subject, of, the agent of, such intentions. Intentions, then, can be identified and individuated independently of anyone who has them. Nor are the first-personal intentions necessarily singular. "We intend . . ." and "We shall . . ." can express intentions of groups, institutions, corporations, governments, when no individual has (or in some cases, can have) such an intention. (Cf. *Zettel*, 48)

Generic intentions can be shared; both parties intend to be the first in the ticket line at the theater. Here their common generic intention cannot be realized for both agents. But if I intend to attend the concert and you intend the same, there is no reason so far why our common intention should not be twice realized. Shared intentions are not the same as plural intentions. In the example just given, I intend to attend and you intend to attend as well. A coordinating relation or coordinating role between us is required for a plural intention, in which case we might intend to attend together. Plural, collective, and conjoint actions are subject to plural intention.

Intentions can be formulated and attributed in a language other than that of the agent who has the intentions. But clearly the languages must have some common structures, referring devices, predicates for action properties and relations, and terms for adverbial or predicate modification. Strong translatability assumptions are made in historical and critical

studies; common sense appears to accept these assumptions as well.

The action predicates and relations and predicate modifiers which appear in the formulations of intentions are the same action universals which appear in action descriptions and in predictions of actions. This part of the data already suggests that "intends" should not be taken as a predicate modifier, operating on action verbs, say, to create new predicates; this would seem to have the result that "felling the tree" and "intentionally felling the tree" or "to fell the tree" and "to intend to fell the tree" are pairs of different predicates in the manner of "is green" and "is pale green" and "is approaching" and "is slowly approaching." Of course, the data so far do not *require* rejection of the predicate modifier view.

Finally, intentions as contents of intending must be the sorts of thing which agents can be said to consider, rehearse in thought and imagination, reason about, adopt, reject, hold more or less firmly, and amend and qualify in one or another way. These contents must also be such as to play a role in the agent's psychology, generate undertakings with approximately co-ordinated or even identical content, and be generated by psychological states and episodes such as desires, beliefs, attitudes and judgments. This suggests that they are going to be in the category of *thoughts*, though other data strongly indicate that they are not thoughts as such; for instance, intentions as such as neither true nor false. Unlike some thoughts and beliefs, intentions do not demand that-clause complements in their formulations.

Various types of theory have been proposed in response to some of these data. We should examine some of these theories.

III

(1) Intentions are merely propositions. This theory is probably motivated by the simplifying assumption that believing or a near-belief attitude (judging, e.g.) is the only propositional attitude suited to a psychological view which avoids notions of causation foreign to science. Intending, as an attitude distinct from desiring, believing or judging, is thought to involve some notion of causation (perhaps, 'agent-causation'); hence, there is a strong motivation in favor of employing believing or a near-belief attitude such as judging something desirable (Cf. Chapter Three, pp. 39–41). Thus, since the appropriate objects of belief are propositions, intentions are merely propositions.

Which propositions are intentions? There is the difficulty that inten-

tions in the ostensive mode contain indexical terms as well as singular pronouns; it is not obvious that the best view of propositions as logical forms has them containing indexicals and pronouns. Even if that difficulty is met, however, the porpositional view of intentions fails to square with a good deal of the data outlined above. For instance, intentions are neither true nor false. Further, simple propsoitions involve predications of properties to individuals. Thus, those which are supposed to be intentions had better be something like "I will A", "We will A", where 'A' is an actional predicate. But an agent may accept such a proposition while not intending and even intending not to A. "I must give up under threats to my family, but I don't intend to." "He knew that he would have eventually to comply with the terrorists' demands; when he did so, however, it was not voluntarily." The situations these describe show that some propositions predicating future action of the agent are accepted by him but not intentions of his. It is not clear why this should not be true of any candidate propositions proposed on the propositional theory of intentions. (Cf. in addition the data Castañeda cites in rejecting the propositional view; see his *Thinking and Doing*, Chapter III. In light of Castañeda's results I shall not spill further ink over the propositional view of intentions.)

(2) Intentions are propositions with a difference; they contain a special type of predicate or special category of predicate. These special predicates have the feature that they are satisfiable by future contingent states-of-affairs only. Consequently propositions with ineliminable predicates of this category exhibit truth-gaps; they are neither true nor false until the worlds to which they are temporally indexed are parts of the actual world. The semantical details underlying such a view are likely to involve complexity. The idea is, however, clear enough. These special predicates look like "A-ing by . . . at t". Simple actional predicates in attributions of present and past action (e.g., "He is A-ing", and "He A-ed") are augmented in intentions and are actually n-adic relations of the form "A-ing, by . . ., at . . .", and the values of the temporal feature of these relations are exclusively *future* times, moments, or stretches of future time. Add, then, the view that future contingent propositions exhibit truth gaps, and the idea of the present view is reasonably clear.

An intention will now be something like "A-ing at t holds of me", or so as to indicate the predicate category, (A*tI), or "I satisfy A*t"; or, perhaps "I satisfy t A*". This differs from "I will satisfy 'A' at t" since the last is a proposition, now true or false, in case I have (or lack timelessly) the property of A-ing at t. But "A*tI" is only true or false at t. "Intention-

al futurity" is treated as a kind of predicate modifier yielding intentions by operations on ordinary descriptive action predicates. "A* at . . . by . . ." is an 'intention function' which becomes an intention type in case (Et) (Ex) (A*tx), a generic intention in case (A*tI), and an intention in the ostensive mode in case (A* *now* I). Intention types are contemplated, entertained and used in considering alternative plans; generic intentions are adopted as future or as further intentions; intentions in the ostensive mode are contents of intendings to undertake or do a thing here, now. Intuitively these contents are respectively: *someone in the future satisfies A (perhaps me), someone in the future satisfies A (namely me), and I satisfy A, now.*

In the data advanced in Section II, I claimed that the very same action predicates which appear in predictions and descriptions of action are those appearing in intentions. On the present view, this would not be the case. The data could be accommodated, perhaps, with the aid of a notion of predicate implication, so that the special 'intentional predicates' of intentions will imply the actional predicates. If "A*tI" is satisfied at t, then at t "A-ed". Thus for an intentional action predicate of the form 'P*', there is an action predicate 'P' and 'P*' results from 'P' by means of intentional modification, represented by '*t'. This, however, will have peculiar redundancies when the action property represented by 'P . . .' is an essentially intentional action property, e.g., signing the contract, looking up the house number, etc. An intention with respect to an essentially intentional action of some type, A, will appear as an intention to intentionally A. Furthermore, one who predicts and agent's intentional action will not strictly speaking employ the same action predicates in the prediction as those present in the intention; though, as we saw, the last will imply the first, the converse is not true. This datum seems hard to accommodate on the predicate modifier view. Nor is it clear how the complications needed for a finer grained analysis of the structure of intentions would be incorporated into a predicate modification view. This is, if we consider the distinctions between intentional subjectivity, intentional activity, intentional manner, further intention, and intentional objectivity (Cf. Chapter Six, pp. 86–88), it appears that only intentional activity (and possibly manner) will fit the pattern of the predicate modification account. Intentional objectification with its strong existence assumptions seems hard to deal with here unless, implausibly, one incorporates it into the predicate, creating a relation and a new predicate containing an individual constant. For example, "Kay intends to eat the pie on John's plate" would have the relata (the pie, John's plate) incorpo-

rated in the predicate (to eat). In that case we could not display the implications we noted in the structured data in Chapter Six. Now we surely have entirely too many distinct predicates and category of predicates for this to be a reasonably workable and suitably simple theory of intentions.

(3) Let us revert to one of the simpler theories. Intentions are nothing but action universals themselves. "Name some intentions!" This theory answers straightforwardly: "to eat", "to greet", "to meet", "to say . . . to . . . when . . . at . . .", etc. Now as the last on this list indicates,the action predicates are noted for their so-called variable polyadicity. The form of an intention cannot be given by merely citing the action universal named by an action predicate; the form can be more or less arbitrarily altered introducing new argument places. Accordingly, this theory should modify its simple straightforwardness with some provision for specifying the 'order' of the predicates which name intentions. Some of the simple intention implications will now become matters which are not displayed by the form of the intention in anything like 'logical form'. Further, expressions of intentions for the future have form and structure which mere names of action universals, even specified as to argument places, do not exhibit. Once again a theory which places all the logical burden on action predicates will have difficulty with intentional objectivity and with the intentionalistic description under which some object is the object of one's intentional action. Of course, on this theory intentions will not have truth values; but the predictive aspect of their expression will not be displayed at all. This theory will place a great burden on pragmatic features and context. Ironically, it is just these features of context and presupposition which remove the slight evidence this theory appears to have in its favor. No doubt *in context* it is possible to reveal one's intention with a simple infinitive phrase.

When in reply to "What do you intend?" I say "to sleep until 9:00" it is understood that I am not merely citing an action property; the pragmatic context gives the reply its understood antecedent, *viz.*, "I intend . . ." If the theory of predicate modification presents complexities which tell against it, the simple action universal theory is simply not complicated enough to deal with much of the data.

(4) In addition to the objective modalities of necessity, possibility and the epistemic modalities, a theory of 'subjective' modalities might be constructed; 'intentionally' might be suggested as a modal operator which takes propositions or sentences into intentions. Thus,the proposition (read "sentence" if one wishes) "I will φ" might yield "I (I will φ)" under this operation, where the initial "I" is the intentional modal oper-

ator, or to avoid this duplication of characters "S (I will φ)", where "S" is the intention operator. This operator takes truth valued contents into intentions, which lack truth values. Since it operates on contents, i.e., contents of psychological or cognitive states or conditions, it is not an 'objective' modality. Since intentions do not have genuine external negations, the negation of an intention will have the form S (I will not-φ). As I am construing this possible theory, an attribution will have the form of S (He will φ). The action predicates embedded in the actional proposition are, of course, the usual predicates as demanded by the data.

One initial weakness of this view is that it fails to exhibit the 'variable range' of intention modification. Since the action predicates and relations themselves have structure and elements in their structure are variously modified by forms of "intention," a whole sentence of proposition modifier does not capture these features. For some purposes, then, this theory is not yet fine grained enough. And since the embedded proposition or sentence on which the intention operator operates is bereft of all intentional contents this view will recast all further intentions in the role of 'by-relations' or some such. "I intend to X with the further intention to Y" becomes S(I will Y by X-ing). Yet not all fulfilled further intentions are fulfilled *by* the instrumentality and means of the initial intentional action. "In order to", "in order to then be able to", "so as to", "in", and others also express part of further intentions. It is unclear how the sentential modifier view will deal with these. They are not, evidently, to be conjunctive intentions, e.g., where "I intend to go to the party early in order to be able to leave early" would go into S (I will go early . . . and I will leave early). But normal conjunction does not capture the further purpose expressed in "in order to be 'socially' able to . . .". This account seems suited, at most, to simple intentions for the future. It seems, too, that nothing short of the structure of predicates and the nature of the 'propositional' or 'predicational tie' in intentions will be suited for the fine grained distinctions and differences reflected in some of the data. We should cater to as much of this as we can, short of a theory which stands in a one-one correspondence with the data; an area which is its own map is accurate though not sufficiently maplike in being unusable.

(5) There is, then, the view that intentions are propositional-like structures featuring a special tie between actional predicates of the usual sort and subjects. This theory postulates a non-propositional conception of predication. Call this "intentional predication" by means of the "intentional copula". (Cf. Castañeda's 'practitional copula' in *Thinking and Doing*.) Let the normal form of the infinitives of action verbs carry this

import. Thus, "I to φ" predicates φ-ing of me in an essentialy and ineliminable first-person intentional way. "I will φ", by contrast, is simply the future tensed sentence, expressive of the proposition that, say, there is some future time at which it is true that I φ. Since intentions do not have truth values, a form of predication of action properties and relations which does not create propositional funtions and propositions seems required by the data. The intentional tie or intentional copula view is especially designed for this part of the data; in particular, it is designed to deal with the fact that intentions contain parts which are and parts which are not propositions. Thus, "I shall tell him your news as soon as he arrives if it is before I leave" is an intention expression with a condition expressed by the proposition "It (then) is before I leave". But intentions can also contain singular propositions with the agent himself as subject; here, then, is where the intentional copula view comes into its own, for it can discriminate the crucial difference between, for instance, "I shall leave if I do not turn on the heat", "I shall turn on the heat unless I leave", and "I shall leave or else I shall turn on the heat". The last is a disjunctive intention with two occurrences of the intentional copula; the former two sentences express conditional (or disjunctive) intentions with one occurrence of the intentional copula and one occurrence of property predication. Given this motivation for the postulation of (or theoretical discovery and formulation of) the intentional copula, one might suppose that additional types of 'tie' might be equally motivated; H-N. Castañeda has proposed some eight or ten different copulas, e.g., intentional, prescriptional, fictional, and ones for contingent and necessary identity, and others. And one might then suppose that the theory has got more complex than the data; do not multiply ties beyond necessity! However, since the intentional copula is, in fact, distinct from these forms of predication we can examine it independently.

The intentional copula account admirably accounts for a good deal of the central data concerning simple intentions for the future. It accomplishes this by its strong focus on practical thinking and the essentially first-personal form of intentions, the conclusions or decisions which terminate practical considerations. It is less adapted to the data concerning the variable forms of intention expressions. This may be because it proposes a 'canonical' form for expressions of intention. So expressions of intention in free standing sentences or in sentence fragments can be formulated 'canonically' as "I to φ". In attributions, the appropriate form is a kind of direct quotation. Thus, if John intends to φ, we attribute the appropriate intention to him directly in "John intends 'I shall φ'".

And employing the indirect reflexive pronoun, we attribute the appropriate intention in a kind of indirect quotation in "John intends that he shall φ". (Cf. Casteñeda's notion of quasi-indicators.) When the presence of the intentional copula is understood, we could emply both "John intends to φ", and "John's intention is to φ" as well as "John intends the following: to φ". Furthermore, the present account of the nature of intentions as contents with typical action predicates and the intentional copula squares with the data concerning intention negation. (But see Chapter Six, pp. 100–104) The differences among intentions, e.g., the generic versus the ostensive mode, will naturally be a matter of the predicates and relations expressed in them. So the intentional copula view will require augmentation from an account of actional predicates if it is to deal with the wider data. Additionally, the generation of intentions from antecedent desires, beliefs, moods, emotions, traits and, where present, practical considering and reasoning, is natural and intuitive on this view. Gerundial forms for expressing these antecedents makes this clear. In an unrealistically simple case we would get John's desiring to φ (with appropriate additional elements and circumstances, etc.) yielding John's intending to φ. In short, the intentional copula will be accepted as a working hypothesis for the remainder of this study. But it must be noted that the need for more complexity and fine grained distinctions will now be the work, not of the copula thoery, but of the account of action predicates and relations which must be added to the copula account. By hypothesis, then, intentions are to be taken as structured contents, contents of the conative states of active agents, which feature first-person reference, the intentional tie between the agent and action predicates, and complex or simple action predicates and relations, sufficiently complex to account for generic and ostensive mode intendings.

IV

Action predicates and relations in intentions have at least the complexity of such predicates and relations as we employ in the attribution of action and the description of the actions and activities of others. We describe what another has done, is doing, will do; how others will do, have done or are doing what they do; what further they have done, are or will do in doing what they do; we describe the manner and the means of their performances; and we attribute possible actions to them in contrast to what they actually do. We employ adverbial modifications of various sorts, including adverbial prepositional phrases expressing when, by what

means, where, in what other connections, etc., they do various types of thing. We credit them with successes to various degrees, complete and partial in connection with both what they do and how they do it. And we attribute attempts and tryings in the sense of actual undertakings. Anything an agent can be said to do in these ways he *can* be said to intend.

An initial difference in category of action predicates has been noted before: the distinction between those things a person cannot both do but not intend and those things a person can do but not have intended. The first are what Anscombe called essentially intentional action verbs, the second are non-essentially intentional action verbs. The second category goes especially with our notion of primary intentional action of active agents, though it includes types of action beyond ones exhibited in mere physical power and capacity. It includes pushing, pulling, putting, arranging, letting go, taking hold of, as well as telling or informing someone, writing or inscribing, marking, leaving, coming and going, insulting, annoying someone, and expressing something. So it includes so-called 'physical-motor' activities as well as 'psychological' activity verbs. The essentially intentional verbs require various conventions or rules and complex behavior patterns with conventional signifincance, including the developments of technologies and instrumentalities of a culture, e.g., paying a debt, buying and selling, contracting, telephoning someone, signaling in code, starting a race by signal, and the like. There are borderline cases at the edges of these types of cases.

Since the essentially intentional actions have descriptions in the form of gerunds and infinitives they are appropriate candidates for the role of non-personal intentions, noted among the data above. Describing a part of a mechanism as *a brake* indicates that the artificer or designer was engaged in essentially intentional activity in its construction or design.

The crucial data, however, concern the "variable range of intention" in fully and less fully intention attributions and in explanations of actions. The so-called variable polyadicity of action verbs seems designed to accommodate this phenomenon. For many such verbs there is implicit room for adverbial modifications of manner, place, time, technique, further purpose and others. Adverbial forms of "intend" also occupy these roles. If 'G' is an action predicate subject to manner modification, 'intentionally' may take such a role; it can also modify the adverbial manner itself. Indeed, many of the manner adverbs, those indicating care, caution, skill, close attention, and so on *are* themselves intention implying. Inadvertent, accidental or involuntary heed is not possible. Just as there are essentially intentional verbs of action, there are essentially

intentional adverbs. Both are understood to be in the scope of an intention modifier in any normal occurrence. The non-essentially intentional action verbs and adverbs get marked with forms of intention modification. In considering what to do an agent may focus attention and thought on an alternative he thinks he has and adopt the intention to do a certain thing with consideration or thought to the manner in which he will do it; there will be some manner or other in which he does what he does intentionally. But he may decide to walk to the corner newsstand and happen to do so slowly; he intended to walk to the newsstand and he merely happened to walk slowly. Here the manner of action is not considered or intend*ed*. Other forms of adverbial modification, i.e., relations of place, time, circumstance, degree, etc., require even mre explicit marking than do adverbs of manner. For instance, it will not be true that he intentionally φ-ed simultaneously with some distant event's occurrence unless he was aware of the event and its occurrence was part of his considerations and plans (reasonably or not does not matter). Typically, limiting adverbs must be marked if they are in the range of intentional modification, e.g., "Intentionally he hardly moved" and "Intentionally his speech was barely audible."

The complex of intentionality, i.e., intentionality modifications of activity, manner, objects under agent's description, and further intentionality, is conceived as a complex polyadic relation. To illustrate once again: Smith intends to heat the metal to just its melting point in order to pour it in these moulds for shot for his antique Colt 44. If this expresses a fully intentional further intention, then Smith (in whose self-reference the name "Smith" needn't appear) intends to do what he understands to be heating of what he takes to be a metal and he intends to thereby melt this metal but only to just the melting point since he understands that the melted metal will pour into moulds at that temperature and he intends to pour it into some moulds for shot casting, which he understands are shot for what he understands is his antique Colt 44. Of course, there are other readings of the beliefs, thoughts, undertakings and so on that go with the target example and show it to be a fully intentional further intention. The point to illustrate is that full intentionality requires some such background, since full intentionality requires (i) intentionality activity, expressed in the action verbs modified by a form of "intends", (ii) intentional manner, means, place, time, etc., as expressed by intentional or cognitive modifications of the adverbs and adverbial prepositional phrases (explicit or understood), (iii) intentional objectivity, as expressed by the agent's thinking of the object of his action in a certain manner,

or under a certain description, *which figures in* his reasoning, decision, etc. concerning the action in prospect, and (iv), if relevant, a fully intentional further intention, i.e., another sequence of verb, adverbial modifiers, objects, etc., which meet the same conditions as (i) thru (iii).

An intention under consideration, as an object of speculation, imagination, and rehearsal, is the predication of such complex relations to some (or any) agent. An agent considers the intention as something one could adopt as part of a plan of action. An intention of the generic sort when adopted contains the same complex relation or modified predicates now self-predicated by the agent. An agent adopts the intention as something to be fulfilled by him at the appropriate future time. An intention of the ostensive-demonstrative sort is simply an intention whose general modifications as to time (and perhaps, place) has been reduced to the present; now (and perhaps, here) is the time for undertaking action.

What must be emphasized, however, is that these contents are only as determinate, specific, or fine grained as the psychological processes of the practical thinking of the agent on the occasion. Mostly intentional actions are spontaneous, habitual or unrehearsed exercises of the agent's powers. Here cognitive and sophisticated conative antecedents are not the focus of practical considerations-what-to-do. If specific contentual antecedents are required to explain or to justify the rationality of the agent's actions, the agent may express them only later, as responses which now make it clear what the agent was doing. If observers attempt to explain and understand the agent, they may attribute specific contents to the degree required to grasp the rationality (or lack of it) in his activities. Observers attempt to understand either what the agent might have or what the agent must have been doing, in light of their application of the scheme of intentional action explanation. While specific, consciously articulated full intentions play a role in many actions and activities, this role should not be exaggerated in an attempt to make all actions fit a determinate pattern. Sophisticated contents are required and cited or avowed in specific contexts in which there is room for them in the application of the intentional scheme of explanation. Interrupt a competent agent in his routine but heedful activities, say in his driving an automobile or his adding figures in a bank book, and ask for details about his intentions; one will get vague general accounts, normally about a further intention somewhat removed from the undertakings he can be seen to perform. He is engaged in intentional activity, of course, since he is not surprised that he is doing, accomplishing, what he *is* doing; pressed for psychological details, he will supply the means and manner of the kind of action in

which he is engaged. These are the intentional content of his activity, in this sort of typical case. Fortunately, we are not the sorts of agents which must calculate and reassess the progress of our actions as means to remoter purposes. Only sometimes is deliberation and detailed practical reasoning needed and present. In such cases, specific contentual desires, beliefs and intentions are required for planning, undertaking, and successful performances where the world cooperates with our concerns.

V

Where practical thinking is fully engaged and active, of course, we require cognitive and conative contents in the form of propositions and intentions. The last are forms of predication of action properties and relations. In practical consideration and thinking they are essentially first-personal. (Whether this requirement implies the basic ontology contains 'selves', 'substantial egos', etc. is a further question. It does imply that we are self-referring agents; that we think of ourselves as such). The predicates and relations in intentions are themselves subject to adverbial intentional modification. Intentional objectivity and its strong existence assumptions require that agents have a range of descriptive predicates and relations in terms of which they identify and think of the actual objects which figure in their actions and activities.

In short, intentionality is present in at least three loci in full intentions: the predicational tie or intentional copula which connects the agent and the actional predicates and objects, the intentional modifications of the action predicates, and the descriptions under which the agent thinks of the objects of his intentions. Added loci of the same sorts are present in any further intentions the agent contemplates, considers for practical purposes, and adopts. In undertakings, these intentional loci provide the 'vectors' for the agent's power, force, or energy as he acts upon and in the environment. Nothing less complex will accommodate the data.

CHAPTER EIGHT

DYNAMICS OF INTENTIONAL ACTION

> A man . . . is acted on at once by an impulse to avenge an affront, by a bodily want, by a call to duty, by a fear . . . He presents them to himself as influences by which he is consciously affected but which are not he, and with none of which he yet identifies himself . . . none is yet *his* object . . . when the man's relation to these influences is altered by his identifying himself with one . . . This is to *will* . . . But in the act of will the man does not cease to desire . . . The object is one which for the time the man identifies with himself, so that in being determined by it he is consciously determined by himself . . . it is not from them, but from himself as affected by them, that the action proceeds.
>
> T.H. Green

I

Intentions appear, typically, in the conscious thinking of agents as the result of considerations, choices, decisions, desires, interests, felt obligations, roles in institutions and groups, emotions, moods, motives, and other traits of character, and other intentions. Intendings are the dispositions, states or episodes whose contents are intentions. An intending is a thinking what to do oneself and, typically, how to do it. The contents of intendings are expressed in full sentences in the typical case. They are also expressed in 'free standing' infinitives, gerundial forms, or appropriate verbs. The form of expression is not crucial to the content expressed, since there are intentions which cannot be expressed linguistically (by children, for instance, and by the inarticulate). Further, some intentions are never formulated by agents who are under the pressure of necessity for action. Necessities of time and contingencies and opportunities do not wait upon linguistic formulations of what one is about to do, is doing, or has done, intentionally. Sometimes, of course, intentions are expressed in explicit ways, e.g., "I shall . . .", "I will . . .", or in free-standing sentences. In these cases, intention contents are expressed in verbs of action, gerunds, infinitives, and that-clauses. Only preferences of style select among these forms. Thus, propositional that-clauses are not essen-

tial; nor are expressions in infinitives essential. If this is correct, appeals to aspects of syntax will not be critical for an account of the forms of intention. Of course, any such account must permit the types of expression there are.

Intentions are both expressed and attributed. for instance, in express attributions, such forms as "He (she) will . . .", and "They will . . ." are typically employed. Forms of the verb "to intend" are also employed. Typically, intentions are formulated in explaining one's action, clarifying one's activities, as well as in getting clear about what to do and in giving others to understand what one will do oneself. And intentions are formulated when attributed to agents by others in their attempts to explain and understand behavior. There is, of course, a variety of things others might want to explain or understand when confronted with the agent's deeds and the agent's practical thinking.

We shall see that both the effectiveness of our practical thinking in planning what to do and the success of explanations and clarifications of the actions of others depend on the presence of content in expressed and attributed intentions. The rational and reasonable character of action and its explanation as unreasonable are only possible through consideration of its intentional content. Yet, content is not necessarily sentential or propositional. If this is so, we can lean on the sentential content of intentions only up to a point in seeking an account of the principles of rational consideration there might be and in formulating any principles of explanation of intentional action there are. If agents were pure rationalizers, matters might be otherwise; if the contents of all action relevant states were purely sentential or propositional, matters might be otherwise. One such difference would be the simplicity of rational psychology.

Further, intendings and other action relevant conditions involve semantic aspects and references in particular. But not all such states make or purport to make reference to objective things and states of affairs in the way perceptual beliefs do. If agents were objective reference machines, causally or otherwise linked to the actual world, matters of assessment and explanation of action would be rather different from what the data suggest. Both the problem of reference and the issues over sententiality pose special tasks for understanding an agent's intending and its role in the dynamics of his actions.

Fortunately, some intendings have explicit sentential, structured contents or intentions, appear more or less transparently in conscious processes and are stated with clear reference. Here we may be 'inscriptionalists' in our account of contents. It is therefore possible to begin considera-

tion with these cases, with the warning that they are not the only and, very likely, not the predominant sorts of cases. As we move away from these simpler and clearer cases, we rely more and more on circumstances and surroundings of agents and on our picture of their special perspective. We rely on our sense of plausibility and on the intricate connectedness of all the agent's action relevant states and dispositions. In the extreme, we are required to decipher his practical thinking in a language in which the contents of his thinking do not appear. But issues in the epistemology of mind and in the theory of action could hardly be expected not to merge sooner or later.

II

Cognitive, affective, and conative antecedents of action, together with circumstances and perspective, go to produce and, when cited in appropriate ways, explain action. More particularly, the presence of intendings among such antecedents go to explain what is or was done and, typically, why it is or was undertaken, maintained, guided and terminated. These remarks represent a high idealization.

For one thing, even in the conception of this unrealistic picture of intentional performance, we presuppose the background of agreement and practice upon which agents and spectators rely in attributing the complex of psychological antecedents and in citing it in explanations. Such attributions must have a plausibility in the view of those who attribute them, and part at least of this plausibility depends upon the background psychological picture of agents which is, in turn, relative to historical, cultural and social settings. Beliefs, desires, intentions and actions form units and not every possible cluster will be found plausible. We demand that we be able to see or appreciate how *that* complex of belief and desire in *that* situation yields *that* intending and how the content of that intending yields the undertaking we identify as the agent's action or as a part of his action. Intentions play a central role in this regard since among conative contents they are exactly fit for realization. Intendings are identified by their content, the intention which the conation of intending embodies.

Yet another departure from the idealized scheme arises from the following consideration. If, roughly, the content of an intending, with its aspects of *what* and *how*, is duplicated or instantiated as the (or a) leading description of the content of the undertaking, the action is classified and partialy explained as *that* intentional action. Undertakings, whether bodily movements or not, seldom replicate intentions in every respect.

For instance, an intention for the future is descriptively generic or general. A rough approximation of its form is a quantified expression in an opaque position. One intends some event to have a certain feature in a certain manner of acquisition. Or, one intends that the object one identifies in a certain way should change (or remain unchanged) in a certain respect. Actual objects of future intentions are themselves, of course, determinate, but as the subjects of intentions much of their determinate character is irrelevant, or ignored, or beyond representation, or unknown. It suffices that we manage to refer or represent them in some manner. As well, the actional predicates in intentions are 'universals', typically expressed by infinitives. Successful undertakings will exemplify such 'universals' and have, as well, a multiplicity of other features in the actual world. What is selected for explanation or as a target for clarification is a feature of the action which putatively expresses what it is in a sense of *identity* fit for actions and a feature concerning how it was generated or produced by the agent's undertaking. If these two characteristics are also in the intention's content, then the complex of desire, belief and intention is, so far, supposed to explain the intentional action in question. But whether these constraints are met in actual cases is often a defeasible matter, liable to annulment on further and deeper inquiry.

Similar connections must hold within the antecedents of action. We require to see how that set of beliefs, desires, emotions, moods, traits, etc. yields that intention. And we accept that there are reasonably clear cases in which we understand how an agent in certain circumstances would intend what he does, given the beliefs and desires and so on he evidently has. His avowals and past performances will naturally play a leading role in such cases. Given that he can do various things, has a certain repertoire of actions available to him, is knowledgable in matters of the ways and means of things, and wants, or values a certain thing sufficiently, and believes that the way for him to secure the desired (or desirable) upshot is by undertaking to φ, given all this, it is evident that he entertain and endorse the intention expressed by him in "I shall φ". If, then, he undertakes to φ, directly by φ-ing or by ψ-ing in order to φ, we feel assured that we have an account of his intentionally φ-ing (or ψ-ing). We have, in sum, a scheme of explanation and understanding fit for more or less reasonable agents in a more or less orderly environment, understood against the background of our own beliefs and perspective.

What we often encounter, outside sparsely described examples in expositions of the scheme, are one or another degree of departure from the idealization. What I wish to emphasize is that departures from the

idealization are recognized pieces of intentional activity because we understand them as departures from the idealized scheme just now sketched. We understand less than maximally rational agents in less than fully orderly environments by contrast with and by degree of departure from this scheme of explanation. There is a partial analogy with our practices in other areas, e.g., in connection with rough and ready measurements in practical situations. We accept and use measurements, of lengths, distances, areas of land and so on which we now know to be actually inaccurate. Prior to the time of the discovery of the facts which show them to be inaccurate our practices of practical measuring and judging were perfectly acceptable; for most of our purposes they remain so after the discovery of these facts. Prior to these discoveries we did not even have the idea of "in principle perfectly accurate measurement." The "in principle" perfection of techniques of measurement is no cause for alarm concerning our judgments and measurements in the actual world. Still, we can better understand our practical techniques by contrast with 'in principle' better (or even perfect) ones. Similarly, the departures from the idealized scheme of explanation of rational action are not suspect schema. No 'in principle' perfectability is a consideration in support of skepticism concerning our understanding of the actions of others. Rather, it is a recognition that there can always be considerations, points of view, deeper appreciations which are relevant to those cases of action which are practically important enough to call for such measures in attempts to understand ourselves and others in action. Let us turn to a description of some types of departures in the dynamics and explantion of actions.

III

In the idealized scheme of action explanation the elements of interest were (a) circumstances of the agent, (b) background beliefs and attitudes of the agent, (c) desires and dispositions, (d) traits and emotions of the agent, and (e) the contents of the above by virtue of which they tend to yield intendings, and, finally, (f) the contents of intendings and the action repertoires of agents which match so that the former can yield undertakings which are exercises of the latter. These are the 'bare bones' elements of a typical example in which a human performance, and action (to φ) or an activity (φ-ing), is to be explained. Departures from the scheme can occur to one degree or another among one or more of these elements.

One of the most theoretically interesting and least appreciated fea-

tures or sets of features of the scheme concerns (f), the generation of intendings leading to 'matching' performance. A crucial datum must be accommodated in accounts of such generation and 'matching'. An agent must be able to consider, mull over, imaginatively test, etc. an intention. He must be able to practically consider *his doing* φ, 'imagine' or think about its being or becoming the case that he φ's, that the world is such that φ, and so on, without his actually undertaking to φ. The needed condition can be expressed as the necessity that generic intentions can be entertained without their tendency to become realized being activated. It is tempting to think that this condition could be satisfied just when the agent is considering certain possible future states-of-affairs. Rather than having to consider the possibility of his intentionally φ-ing, is it not sufficient that he can think, contemplate and consider the future as containing the fact or state-of-affairs expressed by "φ" or "it is the case that φ" or whatever the correct form of this fact would be?

This suggestion does not seem to capture the relevant difference. There is a difference between the agent's consideration of possible future states as such and his consideration of possible future states that he can or thinks he can have a hand in producing. (Cf. Castañeda's 'Future Zone of Indeterminancy' and 'Future Framework'.) The crucial difference seems to be that some possible future states or conditions are 'indexed' to the agent as states or conditions *he* can (or thinks he can) effect; others are only ones he and others can contemplate and, should they materialize, note and observe, perceive and record, etc. Generic intentions are fit for the former role in practical thinking since they are, by virtue of their generality, not what we shall call "intentions in the ostensive or demonstrative mode". These last are what sometimes are called "volitions". Here is a scenario which illustrates the difference. Our agent wishes to eat an apple, believing there to be some in the next room, in the fruit bowl on the table. He intends to eat an apple from the fruit bowl in the next room and proceeds to get up from his desk and go fetch one. But which apple? There are several. Exactly which movements will be undertaken in fulfilling this intention. There are a multitude of ways he can go and many different manners for each? Etc. Consider just his generic intention to eat an apple from the table in a moment or two. Consider him now proceeding to the next room, spying the fruit bowl, seeing a fine specimen of an apple on top of the pile of fruit. His generic intention is now 'reduced', by the changes in perception and time, to the intention expressed by "I'll eat *this* one" (an intention in the ostensive mode).

Generic intentions can guide activities, e.g., going to the next room on

the lookout for an apple; ostensive intentions, which are more event-like, have direct action producing potential. Nothing intervening, he will undertake to eat *this* apple *now*. Accordingly, in considering what to do, contemplating alternative intentions and formulating plans, and in imagining oneself doing something in the future, generic intentions are the contents of such states. Of course, one is typically also considering many other future contingencies, also essentially general, future tensed propositions. Generic intentions can be entertained and adopted. As entertained they help formulate plans of action. As adopted they guide planning and generate, under perceptual awareness, intentions in the ostensive mode. Intentions in the ostensive mode, by contrast, directly produce undertakings, which if successful, fulfill the content of the intention and thus satisfy the generic intention in its original or amended version (in case it was altered in the course of action, due to intervening events, or changes in one's desires and interests, and so on). Generic intentions are generated by the other elements in the explanatory scheme, e.g., desires, beliefs, and action capabilities, etc. They are adopted if deliberation is ended with decision, choice, resolution (in the idealized cases); and they yield corresponding intentions in the ostensive mode under the influence of perception and awareness that *here* and *now* is the time for undertaking action. In fact, there can be a way in which even generic intentions admit of degree; so to speak, their *intentions* may increase under the influence of choices and decisions so that their *extensions* narrow and thus, their satisfaction conditions are more specific. One considers having some apple or other; one decides to have some apple in the house; one resolves to have an apple from the fruit bowl; etc. These are among the dynamics of practical thinking, leading to intentional agency.

One type of departure in the process leading from consideration to intending in light of circumstances, beliefs, desires and capacities of a particular agent with his determinate traits, moods, and values can occur in the generation of generic intentions. Imagine situations which begin in rational consideration of alternatives which produce imagined courses of action. One generic intention generated from deliberation may be to φ at some time or other. One's thought might be expressed by "It would be a good thing if, when Q, I φ". Subsequently, one acts with the intention "I shall φ", but in the interim there is no coherent practical thinking, no consideration of the pros and cons of φ-ing, no imagined scenarios in which one does not φ, and no consideration of whether or not Q, and so on. We have, as it were, two ends of a process with no apparent connecting stages; order at each end of the procedure, disorder or lack of proce-

dure in the interim. A good deal of intentional action appears to exhibit this trait. The lack of closely tied contents in the interim does not render such action unintentional or non-intentional. Some of it may be of the primary intentional form, requiring no specific contentual mental and conative states as antecedents. The content of the learnt capacity of the agent is required here. Other cases may be explained just in so far as the agent can supply thoughts and considerations which appear only after the performance they go to explain. The agent's initial considerations, which did yield "It would be a good thing if, when Q, I φ", are a 'seed' which is connected with his later fulfillment of "I shall φ"; but only after the earlier consideration has borne the fruit of action can he (does he) sketch the intermediate stages. As a matter of fact we will often accept that he intentionally φ-ed because he had previously considered that φ-ing would be a good thing to do; definite stages of decision, further deliberation, and the like are not required. The remaining constraint seems to be that, concerning φ-ing, we must find it plausible when tested against our knowledge of *him*, that "He would think that" and "He would do that". Finally, still other examples of this type may not require the agent's later narrative of what and how he came to φ. A reasonably strong belief that we understand the agent, his proclivities, interests and traits, can support our judgment that "He is the sort who would think and act thusly". Is this the 'hermeneutic circle'? (Wittgenstein's remarks seem apt in connection with cases of the sort so far noted: "The question is not one of explaining a language-game by means of our experiences, but of noting a language-game." *PI*, 655.)

There are many additional cases of departures from the rationalistic scheme of explanation which are for all that types of explanation of intentional action. The schematic representation of the ideal form of explanation is something like: Agent (A), in circumstances (c), with desires (d) and beliefs (b) undertakes (undertook) action (*a*) because of his intention to φ. Thus, A *a*-ed because he had, given c, d, and b. Or, another version of the schematic, because in c, A had d and b, he had φ and thus *a*-ed. We noted that departures can occur at any element in the schematic. Let us now consider some possibilities.

We can imagine cases in which A does not undertake *a*, in spit of A's having φ, d, and b in circumstances c. There are assumptions about agents built into the standard, rationalistic scheme; a short way of saying what they are is that nothing about the agent (A) blocks his *a*-ing, given φ and d and b in c. This is a very idealized picture of agents. It does not seem a completable task to categorize or list the types of facts about agents

which can undermine the straightforward application of the scheme of rational, intentional action, For instance, suppose a mood (m) had by A in c (given that d and b generate the intention φ) blocks A's undertaking *a*. A mood, not as disruptive as depression, can be just enough to undermine the practical resolve normally associated with φ, and , perhaps, with A's desires (d) as well, so that A does not undertake *a*. Yet there may be no trace of m apart from its 'demoralizing' effect on A's intention to *a*. Perhaps we adopt some 'in principle' form of physicalism concerning mood states, but that will not aid in the epistemological project of explaining A's not undertaking *a*. We know how this can happen from our own case. We can imaginatively attribute such a deactivating and interfering mood to others when they fail to fulfill a resolve or fail to act on their interests and desires when given the opportunity. This is not, of course, a simple inductive analogy, since it requires that we know much about the agent whom we take to be similar to ourselves in the relevant respect. And it seems hopeless to expand the scheme in a non-circular way, resulting in something like, A *a*'s in c, given φ because of d and b, unless m or some such. The mood (m) may not be sufficiently separable from A's not *a*-ing; the moods are variable, with unequal effects on thinking and undertaking, determinate just in particular cases; and there are many versions of near-m moods and feelings. We seem to have particular case explanations backed with little else than similarities and imagination and some degree of insight into the particular effect of a mood on another particular agent. We judge and conclude about such cases with expressions such as "A could have been counted on to *a* had it not been for the fact that he was in such an m-ish mood" and we back this with references to understanding how we, too, have responded in similar cases. (We are, in part, making good some claims about the variety of accounts of akratic action.)

Variations on this general kind of case are also possible. A is in c, had d, b, and is in the mood (m) as before, but unlike the last case, A does not adopt φ. Nor does he undertake *a*. The mood, m, may render A just indecisive enough, just indifferent enough to his own desires and interests, or just depressed enough to regard endeavoring pointless, so that A does not endorse an intention for which he has sufficient reason. Nor does A adopt the contrary intention, say not to φ. Knowing something of A, the effect on him of such moods, we withhold the application of the scheme of explanation for intentional, rational action.

Another variation. Things are much as in the previous case where according to the scheme A has sufficient reason for adopting φ and would

normally, undertake to *a*, except now A adopts φ, and refrains from *a*-ing (rather than merely not *a*-ing, as in the last case). A's action blocking mood causes him to see his interests and desires in such a light that he supposes they shouldn't be satisfied; he does not *reject* them. His is a case of mood induced self-denial. We do not substitute an alternatively instantiated scheme in place of the one in effect, by hypothesis, so as to render A's refraining from *a* itself an instance of the fully rational, intentional action catered to by the scheme. Rather, again, we find the scheme not fully applicable because, as we think, we have some insight into the sort of person A is and the effect *on him* of certain moods (or emotions, too). In others, such moods and emotions may have different consequences.

Here is a variation of the variation. Things are much as in the previous case except A's moods and emotions lead him to see his desires and interests as 'turning to ashes'. So that now the scheme is altered ; in c A retains b, has m, but in place of d, there are now new operative interests or desires. His moods enervate, impair his capacity to act on previously held interests and desires. We naturally withhold the application of the scheme, since one of its elements is lacking.

Similarly, moods, emotions and agitations of one or another sort, can be sufficiently strong so that A no longer sees his alternatives and the future (expressed in b) as before. He may see them differently, so that we do have an alternative scheme, with new b's, and these, together with other operative elements, may lead him to another intention and endeavor. Or else the emotional confusion might be sufficient to render his grasp of alternatives and means-ends connections (the relevant ones) inoperative. So no rational action is forthcoming.

We know, too, that similar agents in similar objective circumstances with similar beliefs and desires may respond differently, with different and sometimes incompatible intentions and correspondingly different undertakings. Particular traits of character and personality may not be sufficient to make them *very* dissimilar agents, yet be sufficient to account for some differences in responses. Agent b, unlike A, is on occasion given to risk taking; that is B's trait. Otherwise A and B are as alike as can be. This trait can play the role of explaining differently generated intentions (in their *what* aspect or *how* aspect or both). The variations on this theme are countless.

None of the departures so far considered have referred to differences in behavioral repertoires. We have assumed, unrealistically, detailed equality between agents (A and B) and in an agent (A) over time. But

one with less that full confidence in a certain ability can be led to exclude or tend to exclude some possible means as things he might undertake in order to advance his further intentions. Our particularized knowledge of capacities and our knowledge of the agent's own view of the extent of his capacity will have a role in determining which means-ends beliefs of his we think are likely to be operative and have a role in determining which range of intentions will seem open to him. What goes for behavioral capacities goes as well for physical limits and opportunities. Indeed, all of the considerations which go to make up the condition expressed by "A *can* φ" will set limits to our application of the scheme of explanation of rational action, since we understand that one might not act or undertake an action either because he can't succeed or can't even try or because he believes or merely conjectures that he can't succeed or can't even try. (Of course, he can try to do what he knows he cannot do in order, for example, to show that he cannot do it.)

These cases have been sketched in very general terms. Each admits of a variety of particularized instances with differences among them which can also put additional constraints on the scheme of explanation standardly supposed to rationalize intentional action. This is a matter of the 'fine grainedness' of the contents of the states and conditions in the scheme, i.e., the desires, beliefs, emotions and moods. Initially it is not clear how much significance these subtler differences of content will make. Certainly in some cases the differences will matter not at all. Exactly *how* an agent refers to an object of desire might not make a difference, so long as he indicated it as desirable in some dimension of desirability. E.g., "I want that to eat" said in the cafeteria line, versus "I want some chipped beef on toast with cream gravy, if you please". There are cases where fine grained reference and perspective make a difference. These seem to be a subset of the cases in which the agent's perspective is indicated by features like stress and emphasis rather that by pure semantic and syntatic features of the sentential expression of the contents of the relevant states, since features like stress and emphasis can make a difference to implicatures if not implications. Consider the content of a desire for fame, as delineated in a fine grained manner for agents A and B. Both desire fame. It motivates much of their thinking and practice and dominates much of their fantasy. Yet the focus of their common desire for fame might differ in significant ways and thus account for differences in their intentions and undertakings. Agent A desires that he should come to be regarded as the leading innovator and the single most successful practioner of a technique for diagnosing a certain hard to detect disease. He

wishes primarily to be known as . . . in such and such circles, to achieve the status in that circle of a . . ., etc. B might desire that he gain fame for the same technique, but the focus of B's desire for fame and glory is directed not to position in a circle of experts – that he will get too, if successful. Rather, B's desire is part of a motivational structure aimed at self-esteem justified by HIS accomplishments. One desire for fame is focused on social position and power, on reputation and prestige. The other desire for fame is focused on self-esteem, personal and individual accomplishment, etc. (Freud can tell us about their respective fathers.) As it were, A desires that he should become *FAMOUS*. B desires that *HE* should become famous. A will not be satisfied if his own accomplishments do not become recognized; B will not be satisfied if he goes unrecognized, but his reward is praise of his *personal* achievement. It can happen that neither A nor B is satisfied when his desires are satisfied. But here the point is that the differences in the way in which the desire for fame is operative in the thinking and doing of A and B can be traced to differences in specifying the 'same' contents. No doubt A and B will do, undertake, plan, and fantasize doing many of the same things, but with varying *adverbial* features. A is quick to publish solid results; B cautious. B is secretive lest others steal *his* thunder; A is open so long as other grant *him* priority. If we know this sort of thing about the desires of A and B, a subtler understanding of their doings is available. In particular, the differences between the activities of these two otherwise similarly ambitious agents are revealed as arising from deeper, less obvious differences in what motivates their thinking and acting. The differences might be revealed in "fine shades of behavior".

In deploying the rational scheme in its full application we have assumed that the elements of the explanans are largely or, at least, somewhat independent; the reasonableness of the agent depends, in part, on the semi-isolation of his beliefs, desires and moods. In actual experience various of the elements can alter, infect and color others. We have seen mood influences in some cases already noted. Belief and desire are also intermingling; and the results are not necessarily irrational beliefs or desires. Much will depend upon the circumstances of the agent and his special relation to things and persons. Consider the ways desire can infect belief. Cases range from irrational wishful thinking, destructive of rationality and objectivity in the infected areas, to special requirements for beliefs and doubts concerning some special matters on the part of some agents. For example, concerning the driver of the yellow Volkswagen, I believe he is probably innocent of a certain crime. The driver of

the yellow Volkswagen is Jones' son. Jones and I share all the evidence upon which I base my view that the driver is probably innocent. Jones believes his son is definitely innocent. I think there is some chance he might be guilty; Jones thinks there is no chance he is guilty. Jones, of course, has desires and interests which I don't share. These desires and interests connect with Jones' belief because of the way in which the possible culprit appears in the references of Jones' psychological states. He appears in other ways in my states. In Jones' states he appears as "my (=Jones') son"; in my states he appears as "the yellow Volkswagen driver". Had the same person appeared in Jones' states as "a person found near the scene of a crime", Jones' desires would not (or need not) have come to bear on his tendencies to believe that *that person* is definitely not guilty of an offense. Jones' interests and desires – arising out of who Jones is – do not attach to his beliefs when the last are identified in such a way as to make Jones and me believers of the same, *viz.*, merely that the driver was at the scene. The way in which each of us refers to the same person opens our other beliefs to influences by our desires, interests, hopes and wishes. What could be a rashly held belief on my part can count for loyalty on the part of Jones. Reference in such contexts can, then, make it natural that desires affect beliefs and, in turn, produce differences in response and actions. Explanations of the action will have to be sensitive to these matters, since they make a difference to the contents of action producing psychological states and conditions.

The foregoing is in danger of misrepresenting the view I advocate in that it has tended to over emphasize the role of the contents of the antecedents of intending and action and of over emphasizing the fine grained delineation of contents. Equal emphasis should be placed on the existence of the variety of intentional action and the variety of explanations, not all requiring articulated contentual antecedents. Psychological conditions vary in the degree of content involvement. Purposive action is sometimes explained by citations of relatively content free and relatively unstructured, unconceptualized antecedents. Cases range from purposive responses of a non-voluntary sort (not reflexes) to purposive primary intentional actions and activities of not yet articulate children and other learners. "He ducked to avoid the oncoming object he glimpsed coming at him at eye-level" expresses the first sort of case in this range. Primary intentional actions and activities were discussed at some length in Chapter Two. Children and other learners engage in such activities before they develop articulate and highly structured perspectives on what they do, can do, might do, will do, and so on. They lack articulate

contents for their performances to express. We regard much of their behavior as expressive of *needs* and only thereby and secondarily of desires.

IV

The notion of bodily movement has played an exaggeratedly prominent role in discussions of action and intentional action. Sometimes it appears to play the role of a 'bare particular' to which different action descriptions are assigned, depending on the agent's antecedent desires and beliefs. In other accounts it appears to play the role of a foundational *given*, from which actions as such are constructed or inferred with the aid of theories of what the movement is a part of, rather like the sense data which were once thought to be the non-logical components of our constructions of material objects. On such a view bodily movements, like sense data, are given; we must determine what actions they 'imply'. Some recent psychologists have taken a more plausible view. Thus, in connection with the actions of animals and in connection with the organization of verbal behavior it has been noted that "the pattern of units at one level [the level of muscle movements, e.g.] can be indicated only by giving the units at the next higher, or more molar, level of description" (G. Miller, E. Galanter, and K. Pribram, *Plans and the Structure of Behavior*, 1960). I wish to support this perspective in connection with the role of 'bodily movement'. I think one should be initially suspicious when it is noted that 'bodily movement' is used in such a way that non-movements or refrainings count as bodily movements.

Views which attach either metaphysical priority to such movements (the bare particular) or epistemological priority to them (the sense data analogy), fail to recognize that the concrete happenings with which inquiry into action begins and ends are performances, i.e., *instantiations* of actual or contrived and constructed action infinitives or action descriptions. These, of course, vary in complexity and in their requirements of application. Some are physical relations whose relata are agents and objects. E.g., A pushed X over, A moved Y from p_1 to p_2, and A arranged (W, U, V and X) in the order (U_1, V_2, W_3, X_4) and so on. The action universals, properties, or relations naturally have some bodily movement(s) as their tokenings. The set of possible tokenings is not specifiable in 'bodily movement' terms alone, since there are an indefinite number of ways the agent's arms, hands, trunk, etc., etc. could move and produce the changes expressed by the action predicates and the relational expressions in these and similar cases. The bodily movements which in fact instantiate them are specifiable only by reference to changes which actually satisfy the predicates and relations when these happenings occur. From the point of view of the spectator of such doings, the bodily

movements in these happenings *are* the tokenings of the predicates and relations of the doings or happenings. From the point of view of the agent of these doings, the bodily movements are, typically and predominately, the spontaneous exercises of his physical ability to *push* something, *move* something from place to place, *arrange* things, etc. Agents can and do calculate, consider, try various alternative ways of producing physical changes of a complex and serial nature; "if I grasp the thread between my left thumb and index and pull the loop gently with the index finger of my right hand this will soon free my left hand to hold the material in place as I attach it to the shank of the hook onto which I am tying this gossamer fine material" etc., etc. These considerations concerning 'bodily movements' are possible only in so far as the agent tying the artificial fish lure or artificial fly has the power and control normally associated with a degree of manual dexterity; such dexterity is not normally a matter of considerations. Dexterity, powers of movement and locomotion and their voluntary control are prerequisites for very many happenings in which we are the agents of change. Consideration of how to use these powers is sometimes relevant, where the changes are complex and ordered; even here practice, training and success typically obviates the need for considerations and thought about 'bodily movement'. The notions of power and control, both native and practiced, in contrast to bodily movement, appear more nearly the fundamentals of voluntary action and primary intentional action. The multitude of capacities and the plasticity of the manner of exercising such powers are the grounds for physical agency. Some 'bodily movement' or other will occur when an agent makes a change in and among the objects of his environment, but the more molar level of voluntary action and primary intentional action has priority since it is by reference to the predicates and relations at this level that we have a line on the bodily changes which express or instantiate these actions and undertakings.

It helps to see the force of the view if we consider cases in which cognitive and conative antecedents play no role (or very little role) in our identification of purposive agency. The relevant science is ethology, the study of animal behavior, which proceeds with a minimum of information about the 'mental' life of the species and individuals under study. We make and confirm certain assumptions about the normal perceptual capacities of the typical member of the species and we attribute a few natural 'needs' to it, based on evolutionary, genetic, and physiological and anatomical information. Then we proceed as follows.

Certain species we study – the elk, the gulls, some fishes, etc. – all

rather far removed from the human species, are known to behave purposively in relation to other animals of the same or different species and in relation to their environments. Whether in the end we decide to characterize their behavior in purposive terms, it is clear that a great deal of animal behavior is goal directed. Various species of fish and birds already extensively studied in the wild and in captivity *engage* in (there is no better term) standardized, even ritualized behavior. Ethologists come to understand and explain the species specific bodily movements of animals, at the species and individual level, by isolating the functions and means-ends connections in the animal's behavior. We isolate stretches of the animal's activities and identify it as a unit, describe it as an ethogram, by reference to typical responses to typical stimuli and by reference to means-ends connections. *Then* we can describe such units in the language of nerves and muscles, study the evolutionary and genetic sources of these units, and compare them with others, first identified in other purposive terms. Prior to purposive identifications, many of which are social and relational, the particular muscular activity involved in an agression movement, for instance, cannot be identified and distinguished from the bodily activity involved in, say, a preening gesture. There is, then, this apparently unavoidable order to our scientific study and understanding of animal behavior. First we observe the animal, armed with some information about its perceptual capacities and some assumptions about needs. Second, observed behaviors and patterns of behavior are described in purposive terms. We form and confirm hypotheses about what the animal will do in certain circumstances, what it will do next, when a certain activity will terminate under stimuli from other animals and the environment. Thereby *types* of behavior are distinguished, e.g., courtship, aggression, denning, food search, stalking, predator avoidance, and so on. Only then are the specific muscular activities of a given behavior isolated, related developmentally to other such activities, and related genetically to other such movements. Here is an apt picture of the role of 'bodily movements' in the life of the individual and species. It articulates the *life* of the creature (as speech articulates the life and thinking of humans.) Movements have their significance by grace of their roles and functions in the life of the animal. Isolated and abstracted, they have no animation.

Intentional actions are normally voluntary movements understood as means to ends. Apart from all natural and conventional means-end connections, bodily events dependent on human beings would be mere movements, bodies in motion. The concept of such a mere movement is,

however, an abstraction. Mere movements are not bedrock components of human actions. Some movements of human bodies are human actions because of the intentionalistic descriptions that count them as means to ends, as handles on the world and recipes for changing. Unlike the targets of the ethologist's study, some human behavior is performed in the guise of thoughtfully intended, goal directed projects. Like the ethologist's targets, however, human actions are identified as specific, intrinsically intentional proceedings (greeting, signaling) or contingently intentional actions (offending, hitting) by virtue of their roles. Intentional actions are identified by reference to their ends or standard results. The same events can be otherwise identified, but they will be identified as human actions in other terms only after they are intentionalistically individuated.

These last points can be illustrated by attending to the range of answers to only two of the questions that are raised in enquiry about simple human events. The questions are: What is he doing (are you doing)? Why is he doing that (are you doing that)? An answer to the second presupposes some sort of answer to the first. The range of appropriate answers to the first question includes the intrinsically intentional action descriptions as well as the contingently intentional action descriptions, e.g., "He is signing a binding contract" as well as "He is interfering with our sleep". There will be no clear target for the question Why if there is *no* answer available for the question What. But some answers to What are also answers to versions of Why. If this is so, then in the appropriate cases, to intentionalistically identify (answer What) is to explain (answer Why). There might always be room for more explanation, in the same or in a different mode of explanation.

Suppose "He is tying the rope to the boulder behind him and tying it around his waist" expresses *what* he is doing intentionally. That is, he is knowingly (not somnambulistically) tying what he knows to be a rope to what he knows to be a boulder situated behind where he plans to sit near the edge of a cliff and he is intentionally tying the rope around his own waist too. The question Why with respect to what he is doing can now be specified in the following ways: Why is *he* tying the rope to the boulder behind him and tying it around himself (in contrast to *your* doing the tying)? Why is he *tying* the rope to the boulder and *tying* it around his waist (in contrast to looping or wrapping it)? Why is he tying the *rope*, etc. (in contrast to the nylon strap)? Why is he tying to the *boulder* (rather than the tree)? Why is he tying the rope to the boulder *and also* to himself (rather than merely to the boulder)? Why is he tying it to the boulder and

also around his *waist* (rather than around his leg, throat, or your arm)?

In doing what he is doing we may presume that he has a further intention or purpose and that he knows (believes) that by doing what he is doing he will also do something further, *viz.*, what answers to his further intention or purpose. He is anchoring himself on the mountain in order to put his climbing partner on belay. We can answer the question Why in some of its specifications by citing what further the agent is doing in doing what he is doing intentionally; what he is doing, in these circumstances, is a means to his further action. It might be the only means available, and if so, citing it would also answer why it is necessary for him to do what he is doing. Other answers to the various specifications of Why might go as follows: "*He* is doing it, because of the two in the climbing party he is the one in position to do it." "He is using a particular tie rather than a loop since it is *called for* here." "He's using a rope; they have no straps." "He is anchoring to the boulder; it's secure and the tree is not really near enough." "He is tying around his waist because that is where the pull on the belay rope will be etc." These possible replies would explain why he is doing the things in question, not that they explain all that can be or always will need to be explained. Nonetheless, given what was asked in the series of Why questions, these answers do explain why he is doing these things by citing *what* he is doing, i.e., by citing his intentions in acting, his further intentions in acting and his knowledge and opinion concerning appropriate means-ends relations in his actions. Hence, intentionalistic identifications of an agent's voluntary movements can explain why he is performing his deeds or why he performed those deeds. We accept such accounts because we know or can come to know the means-ends connections and the in-order-to relations an agent's actions exhibit.

Analogous to the ethologist's understanding of purposive animal behavior, we understand in-culture human actions by identifying *what* agents are doing and thereby answer part or all of the range of Why questions of the sort noted above. Having understood this much about a dated particular event, we may wish to press other questions, in the same or a different mode. Questions in the same mode include: Why is it desirable to use belay here (safety rope)? How capable are these climbers anyway? Is the novice on belay? Etc., Etc. Examples of new questions in a different mode are: What in his brain triggers each sequential muscular unit of movement in this series of events? What is the neurological mechanism which accounts for the perceptual feedback which triggers the next muscular movement in the series of events present in tying a rope knot? And there are many others of similar and of radically different sorts. But

questions in this last mode get their topics in particular cases of actions and activities from the intentionalistic accounts of actions and activities explained in accord with the intentionalistic scheme and its variations.

The notion of *undertaking* has been employed in the above without explanation. Undertaking, endeavoring, trying, setting out to do something, etc. are notions to be met with in typical accounts of the etiology of intentional action. Some theories assign such notions a secondary role in connection with 'basic actions'. Since I have not used the notion of basic action or the notion of actions which are not basic, undertaking or setting out will not have a modest role here. (I have avoided the use of the concept of basic action in part because of its association with attempts to bridge the Cartesian gap between mental and physical events – a gap which needs no filling on the activist view.) By "undertaking" I mean to include, prominently, the *initial* whole unit of action which an agent produces or exhibits in exercising his native and acquired action capacities. An initial unit of action derives its unity from the fact that its purpose is intrinsic; such actions, only typically bodily movements, are essentially means to purposes which are determined by the physical structure of the agent. Examples: turning the head to identify an object on the periphery of the visual field; a distinctive movement or arm, hand, and finger to grasp an object; the distinctive movement of the thumb and index to grasp a feather; the distinctive intake of breath to begin to speak; and similar movements on the part of agents with voluntary control of normal physical structure. The concept of native physical capacity which has appeared in these studies includes these undertakings in its extension.

In developed agents undertakings are distinguished by control and the absence of surprise. This is to say that they are not involuntary, not reflexes; but they often involve the same musculature and other bits of anatomy as involuntary responses and reflex movements. Moderately developed agents 'know' their undertakings, at least the simpler undertakings, without requiring observations of them. There is a complex story to be told here concerning such things as eye-hand coordination, the role of visual cues, kinesthetic sensations, perceptual feedback, servo-control systems, and much else. The outline of the relevant facts are already more or less in hand in physiological psychology. The central point for my purposes is just that the necessary conditions for voluntary control of such undertakings, conditions examined physiologically and neurologically, do not include bodily *observation*. At most a moderately alert state of awareness is necessary. I will not add my speculations to a field requiring detailed physiological study. What is already clear is that children can

be trained as can those who must 'relearn' or 'reacquire' certain physical abilities after traumas of one sort or another. Training employs directives, imperatives, orders, and so on, under the clear assumption that physical undertakings can be voluntary and controlled. Training seeks to remove the agent's surprise concerning his movements, what his limbs do, how his movements coordinate, and so on. After training has accomplished its goal, self-observation becomes a hindrance, even a blockage of ability.

An ability is by its nature something which can be exercised. Physical abilities and their tokenings in undertakings (those undertakings which are physical) provide the agent with the ways and means of primary intentional actions, and, in time, sophisticated intentional acts and activities. Physical undertakings, in this technical sense, are the initial expressions and articulations of the agent's intendings. Types of undertakings are genetic and evolutionary features, largely species specific, and have survival benefits in the standard environment. They are also characterized, in higher species, by plasticity and variability. They are, in short, robust features of successful individuals and species. In higher species they clearly come in units of complex underlying muscular movements and all the supporting physiology and neurology required for complex, serial muscle movements. We undertake to *grasp* objects; the complex and variable musculature and its conditions do not occur singly, 'atomically' as they are in our anatomy. The unit of usable ability is *grasping*.

Trying is not the appropriate general category for the technical notion of *undertaking*. Trying appears to require sophisticated cognitive and other antecedents. As well, trying presuppose advanced behavioral repertoires which themselves presupposes undertakings. Trying is altogether too 'intellectual' a matter for present purposes. In one sense of 'trying' it is an action as such. We can, for instance, desire, decide, intend, choose, etc. to try. We can be ordered to try or to stop trying. Trying, like acting, can come as no surprise to the agent. ("I didn't know I was trying to do *that*!" as a joke.) One engages in trying as in activities. In another use of 'trying' it seems reserved for things which are testing, relative to the agent. An agent in full control of his body cannot be said to try to move; there is nothing testing about his movements. The notion of undertaking is a concept of ability; a 'failure' does not imply the loss, temporarily, of an ability; failures over time carry this implication.

In short, a grasping movement is an undertaking by one who has the use of his hands in the normal way. Undertakings depart from this paradigm depending on the details of the sort of ability and action in question.

Typically and fortunately undertakings automatically activate and animate movements on the part of agents with normal abilities.

V

The dynamics and structure of intentional action have their source in the applicability of the explanatory scheme, with is variations, and in the dynamic, active nature of agents with powers and acquired capabilities. Another condition of the possiblity of intentional actions open to agents has yet to be examined. This is the contribution of their cultural and social setting and inheritance. Beyond their physical capacities which provide for clusters of possible undertakings, clusters of repertoires of sophisticated actions and activities are provided by customs, conventions, and culturally fashioned things to do, expectations of agents, and culturally determined ways and means for fulfilling these expectations. Psychological internalization of these expectations occurs in development and gives the agent added powers of action. As well, it provides the agent with a means of understanding the actions of others. He becomes one of them (*us*) as he learns the standards which are in force in various circumstances, what desires and interests are to count for, and what and how to perform in a variety of circumstances. So he begins to have the materials for the application of the intentional scheme of explanation. His behavioral capacities and his ability to begin to understand others appear to develop as one. The agent comes to share the judgments and practices of his society and wider culture and is in a position, on these bases, to enquire concerning the action of others, and to engage in considerations concerning his own actions. These general but crucial facts are difficult to describe in nonbanalities; what is clear and what needs to be appreciated in spite of its obviousness is that agents are products of physical nature and culture. These are constraints on *any* plausible conception of agency.

As a product of his culture an agent simply does, without justification and considerations, share the practices of others and agree in their judgments. Agreement in judgments and practices consititutes, I think, part of Wittgenstein's elusive "forms of life". Physical capacity provides for action which can be intentional; contingently intentional action-types such as putting something somewhere, pulling, pushing, hitting, etc. The capacity provided by inculturation provides for action which is of the essentially intentional type (greeting, contracting) defined by the conventions in effect (Cf. Anscombe's list of both types, *Intention*, p. 85). Excercises of physical ability can be channeled into the fulfillment of

further intentions and intentions for the future where these last are expressed in the descriptions afforded us by the conventions in force. Included in the *culture* is our science and technology and the artifacts created by technology. Socrates could not have been in a practical psychological state *we* identify as "intending to shut off the microwave oven before mixing the martinis". Nor was Socrates capable of applying the scheme of explanation in the way we can; the psychological technology (e.g. Freud) and literaure of a cultural period gives agents sophisticated categories of beliefs, desire, mood, emotion and the rest that may be foreign to other periods of human history. Sophisticated intentions are not normally fulfilled in the types of action which fulfill intentions such as putting the cup on the shelf. There is no specific type of behavior or movement answering to the sophisticated intentions provided at the most sophisticated levels of culture. What physical actions or activities fulfill intentions which Freudian and other psychologies yield? Their 'fulfillment' consists in their articulation in the thoughts, attitudes, emotions, and psychological responses of the agent who has them, even when the articulation is confused in the thinking and feeling of the agent. Unlike "intending to put *this, there, now*", these sophisticated intendings are fulfilled, not in physical proceedings, but as they color and add to the accounts the agent and others give of his action and the thoughts and feelings which 'surround' actions. Notably, both trained and naturally insightful observers, who have the concepts of such sophisticated matters, do often agree on how to describe the agent's thinking and doing. Insightful observers, too, are 'trained' in our culture to use the vocabulary of subtle and sophisticated attitudes and their contents. Wittgenstein might note: These games, too, are played.

Among sophisticated and culture relative intentions we must count some of those which can characterize an agent's whole life or large stretches of it. These dispositions and proclivities express the agent's standing values, interests, and plans. They clearly operate as constraints on his more specific practical thinkings. But even this may be a function of other traits of his personality and even his power to imagine personal alternatives. Some standing intentions and life plans are sufficiently general so as to make cross cultural comparisons possible. Some of the virtues may be counted among these standing intentions, although the range or extension of some virtue concepts can differ with differences in culture as with differences in social grouping within a single cultural tradition [illegible]larly for vices. What an agent is *like* and *how* he conducts his life a[illegible]ers identified by means of standing intention, e.g. to be

honorable, to be truthful, lawabiding, courageous or legal. Unlike intendings in the ostensive mode, expressed in "I shall φ here and now", many of the standing intentions admit of two types of 'negation'. (The negative of "I shall φ" is just the contrary intention "I shan't φ".) An agent can purpose to ignore truthtelling when it suits him and furthers his other intentions. He may be disposed to conduct himself in that way. He is pernicious or vainglorious or full with calumny and so on, depending on the circumstances and manner of his expression of these traits. Other agents may exhibit lack of virtues in their standing purposes in less venal ways; with a degree of innocence they are unminding of the virtues. Standing intentions, then, can be lacking because of a venal character and can be lacking because of the mere absence of traits which count as virtues.

Besides concepts of virtue and similar concepts of character, other standing aims and purposes are expressed by choices of roles and relations in persisting and in temporary social groupings. And many such roles and relations exist nonvoluntarily. In applying the scheme of explanation to an agent's undertakings it is often relevant to consider what the agent would do, given that we know of his voluntarily, and nonvoluntarily acquired, social roles. Given his beliefs, desires, moods and emotions he *would* intend thus-and-so because of the requirements of his roles and 'station'. Explanations in this style will explain the generation of intentions and actions just because we understand the roles as typically involving certain standing intentions. An agent fails in a certain role if his intentions are not generated or produced in a way which is sensitive to his roles and the expectations associated with them. Evidently we must wield concepts of intentions, such as standing intention, in employing both traits of character and social roles and expectations when we apply the scheme of intentionalistic explanation to particular persons and their particular activities. Rational moral psychology is a substudy of rational psychology only because we have the use of the concepts of standing intentions available in the cultural and social setting whose agents we wish to understand. Unlike intentions which appear in conscious deliberations, standing intentions are not always best understood by those whose intentions they are. Practicing historians know this, for instance, and this fact complicates some of their projects.

VI

To recapitulate, we have a scheme of rational explanation for intentional actions and activities which has its paradigmatic application to conscious considerations of sophisticated agents with the use of their normal, physical and active powers. Citations of beliefs and desires as well as moods, emotions and features of character, perspectives of the agent on his social and personal roles and relations as informed by his culture, all these have a role in generating the intendings which yield his undertakings and guide his activities towards realizations of his further intentions and his intentions for the future.

Departures are normally the rule rather than the exception. (The 'standard' defining example of a species of flora may seldom be found in the wild, with its variations in conditions of growth and other conditions in the environment. Similarly for human performances.) We can understand departures by virtue of our application to one or another version of the rational scheme of explanation. Some such types of departures were described; they are distinguished by the dimensions along which they depart from the scheme. The scheme of explanation is not open to refutation by counter instances. It is not a simple empirical generalization on instances of intentional action, although its application requires our knowledge of many empirical, singular and general propositions, about particular agents and agent-types. The fine grainedness of contents of mental and conative states also plays a role in the degree of departure from the ideal scheme, since the mode of reference or perspective on the contents possessed by the agent may differ from that of those attempting to comprehend his actions and activities.

We have noted, too, the joint conditions on human performances, *viz.*, the possession of native and developed physical capacities and skills, and the contentual *know how* in the ways and means provided by the society and culture. Without the conventions and their expectations in force at a given time we could not be the sorts of sophisticated agents we are.

Undertakings, in the somewhat technical sense introduced above, are the normal causal upshots of intendings. They are the ways we set out to fulfill intentions embodied in intendings. In the typical case they are given by the physical capacity of agents. The dynamics of intentional agency need no special efficient causes or mechanism beyond the causal connections and workings provided by our natural constitution. Practical thinking, intending and their contents channel and animate human per-

formances and their consequences in the environment. Included in the channelized activities of agents is their maintaining environmental 'status quo', when they are *doing* things to forstall changes.

These chapters do not attempt to delineate the wider scientific metaphysical views in which agents and their doings would find a niche. These are prolegomena to such theorizing. It is evident, however, that the wider view will place emphasis on our animality as on our rationality, and thus will emphasize our evolutionary, biological and organic nature and that it will have a likeness to some form of the metaphysics (though not especially the epistemology) of pragmatism. What needs to be remembered is that the study of animal behavior, especially 'in the wild', gives us understanding of both species specific behavior and the individual variations among particular animals. Since we share a long stretch of history in common with them, similar techniques for understanding and explaining ourselves are to be expected.

Underlying the explanatory scheme and supporting its application is a causal nexus or set of causal conditions, in a sense of 'causal' which we must now examine. As should be already evident, however, causation in agency will have some special features, not shared by 'physical' causation. And, the causation of agency and the explanatory scheme of explanation will prove to be mutually dependent.

CHAPTER NINE

AGENCY AND PSYCHOLOGICAL CAUSATION

> . . . only our deeds hold up before us the mirror of our will.
>
> Schopenhauer

I

Agents are physical, living animals. They are the loci of movements and changes. They are also thinking, intending subjects. And they exist and develop in historical, social settings. There are among the facts we must fit together in a coherent picture of agency.

Agents are perceptually responsive and sensitive to the physical and social environment in modes of contact which are direct, causal, referential, *de re*, as well as in more conceptual, mediated, theoretical modalities. But things which are not agents can also be "perceptually" responsive. The definitive feature of agents is, then, their purposive and purposeful actions and stretches of organized actions or activities. Among purposive activities are the goal directed and controlled behavior of animals and the primary intentional activities of human agents in development and in later habitual activities. Purposeful activity is present in undertakings which express and manifest thinking, planning, desiring, conative attitudes in general, and the emotions, moods, traits and beliefs of agents. Purposeful activities express the content of these dispositions, conditions, and episodes. In particular settings and circumstances, purposive and purposeful activity manifests and is, thus, "criterial" for contentual states and dispositions. It is on their basis and, in suitable cases, on the basis of verbal activities, than we can become certain of an agent's contentual dispositions and states. (This does not amount to a "criterial analysis" of contentual dispositions and states, since taken as such it would be circular.) No living animal without the feature of manifesting the contents of its states is an agent in the present sense of "agent". Ducks, grizzly bears, beavers, elks, and so on manifest contentual states, as ethologists know in considerable detail. Humans (and perhaps a few other species) both manifest and *express* contents in very sophisticated ways, including in language. While the notions of manifest-

ing and expressing are degree notions, the expressive feature of human activities does appear to be a species distinctive power. Manifesting and expressing content are capacities of movement, control, and behavior which differerent creatures have to differing extents and in a variety of circumstances, to one degree or another independently of stimulus conditions. Non-peripheral control of physical action is a prerequisite of agency. So, of course, are suitable physical mechanisms and structures. Thus, ability, power, control, skill and the like are crucial ideas in this account of agency. Agents have the power to initiate and control what are contentual, purposive, and, later, purposeful activities.

Implicit here is a sense of "can" which includes the notions of capability and opportunity of doing one thing or another. Over simply, agents *can* do things of, say, type A(movements, undertakings, or action of a certain sort, depending upon their psychological-physical endowment) and, normally, *can* do x and can do y . . .etc., each of which is a tokening of activity of type A. Doing x, of course, may remove the opportunity for doing y, an alternative tokening of A. For a sophisticated example, given one's ability to express in action the intention (content) to, say, raise one's right hand above the top of one's head on signal, one can do it in raising one's hand four inches above the top of one's head and by raising it two inches above the top of one's head. If the agent does the last, the opportunity to do the first is thereby removed.The capacity to have done either can remain when the opportunity for one is removed by the realization of the other. We will be concerned with the "can" of power and control rather than with the "can" of mere opportunity, though in actual cases of power both conditions are normally present.

Anthony Kenny has pointed out that powers or abilities compare and contrast with opportunities in a variety of ways. For instance, abilities are 'internal' to the agent while opportunities are often, at least, matters in some sense 'external' to the agent. Abilities are positive, empowering conditions while opportunities are often the absences of preventing conditions. Different ways of expressing future powers and future opportunities are present (in English) by "will be able" and "can". But the present ability to do a thing includes both the power to do it and the opportunity afforded by the agent's situation (cf. Kenny, *Will, Freedom and Power*, pp. 132 ff.). In the following discussion, we will focus on the idea of power and ability, which is central to the present notion of agency. An emerging thesis is that agency is a fundamental animal-life phenomenon; the idea of *the agent*, cut off from the active life of the organism or animal, is an abstracted, deracinated, version of this notion of agency. Thus con-

ceived, the agent is incapable of animating, directing and embodying the contentual states which, in part, characterize the life of the animal.

In a naturalistic framework, intentionality in the sense of animated, living, directed activity subject to psychological causation goes all through agency and experience. Psychological causation, discussed in more detail below, is present from simple perceptual responses to complex perceptual judgments and to conceptualization in theory construction. In action, psychological causation is present from primary intentional action to sophisticated, purposeful intentional performances. The presence of psychological causation is presupposed by our application of the intentional scheme of explanation. Furthermore, there is no possible justification for the claim that our practices and modes of thinking in employing the intentional scheme of explanation *fit* the natural world we encounter. Rather, this fit is an ultimate but contingent fact. All justification of action and belief depend upon this radical contingency. It is an unjustified ground of justification of action and belief.

What needs explication in this framework includes (1) the notion of purposive power and ability, (2) the notion of purpose*ful* power and ability, (3) the less general notion of the locus of directed energy characteristic of agency, (4) the underlying idea of psychological causation, (5) the notion of agency itself, (6) the conception of "willing" which has played a role in some accounts of action and agency, and (7) the derived notion of "the agent". These are the projects in the remaining sections of this chapter.

II

The idea of power has been suspect since at least the thinking of the eighteenth century. It can be uninformative as hell when used in 'expanations' of the features of various sorts of objects, e.g. the traditional example of the sleep producing feature of opium. In many cases, however, appeals to the powers of an agent are neither circular nor unexplanatory. This is especially evident when one distinguishes the aspects of agency and action which can be the targets of explanation. Briefly, these aspects are, for instance, expressed by questions such as what did he do? why did he do that? what caused him to do it? what was his purpose in doing it? And, how could he do that? how was he able to do it? The last two in the series of queries are often answered by reference to an ability, skill or power of the agent. Skill, power, and ability, in turn, can be explained independently of their exercise in particular cases by scenarios which focus on the

agent's physical "gifts", training, past history of attempts, practice, partial successes, modifications of techniques and successful performances. Underlying training and practice is the physical capacity (physical power) the agent has. The absence of such a power can be cited in explanation of why an agent or a type of agent cannot do certain sorts of thing. *Power* in this sense is not a suspect idea. A main component is physical structure (the opposing thumb of humans, for a positive case, and in the case of the Grizzly Bear's inability to climb trees, its physical bulk and claw structure (contrast the Black Bear)). Physical structures can be the result of training, practice, and the interaction of various physical structures over which the agent has control.

The notion of control, a second element in power, is not widely discussed though often cited in discussions of agency and action. Here the root idea is plasticity and modifiability of action in light of failure, partial success and training in realizing natural, conventional or conscious purposes. Again, lack of physical control of limbs, eye focusing, bodily orientation and so on go to explain the inability of agents and failures of agents to accomplish various tasks they have or which they set themselves. Among the requirements of control is perceptual 'feedback' or perceptual awareness by the agent of its movements and their results and causal consequences. While kinesthetic sensation is, no doubt, a necessary condition for control (as evidenced by its absence in anesthetised limbs) much of the perceptual 'feedback' is, fortunately for the agent, neither observational nor necessarily conscious. Habit requires this. It is not necessary to elaborate the details; they are well known in anatomy and physiology. A moderately specific anatomy and functional physiology is presupposed by the range of power and control types of agents have. Most physical powers are 'wired in' but are modifiable and controlable, species specific movements. Their plasticity consists in the fact that the same musculature can be employed in a number of different tasks and in attaining a number of different native and acquired purposes.

The power to form and adopt purposes in the form of intentions is a considerably more complex capacity. No doubt it requires a complex nervous system and corresponding perceptual capacity. Included in the power of intention formation is the power of intention generation, i.e., the capacity of the organism to form, consider, select for special consideration, adopt intentions, and make decisions because of its perceptions, needs, and in more sophisticated cases, desires, beliefs, emotions, attitudes, promises and so on. In the most sophisticated and least exemplified cases, intention formation and generation can be conscious and deli-

berative. Here is the domain isolated by philosophers under the rubric of "the practical syllogism". But in fact the range of purposive activity is wider than this domain. Many species, including the human, are capable of invention and creation of new methods for fulfilling purposes. The highest level of invention and plasticity of responses to purposes and intentions is, we believe, exhibited in our species. Careful animal studies indicate that there are borderline cases in connection with instrumental intention generation. This is what one should expect from the vantage point of evolutionary theory in species with complex structure.

The concept of agency has, then, the feature of physical power and appropriate modifiability and control of organized behavior, given the species under study. Distinctively for humans, agency includes as well the control afforded by the capacity for intention generation. Human practical thinking is the *loci* of intention adoption where it occurs in such a way as to vector the physical power of the agent. Directed power, guiding intentions, controlled and even calculated responses are in the domain of human agency. Agency, one can say, is the life of the agent; agency is channeled, employed as means to ends, given direction and efficiency by intendings, whose contents are intentions. In what follows I will simplify the picture of agency for purposes of its further exposition. The simplification will aid in bringing to light the embedded concept of psychological causation (or *a* notion of psychological causation) which is essential to this notion of agency.

III

Living agents are *loci* of organized activity. Organization is a matter for psychological causation. In simplified cases of fully intentional action which is more or less successful in fulfilling the agent's intention, psychological causation requires property transference or even property reproduction (in the manner of Cartesian causation). But psychological causation can appear mysterious when melded with the view of agents as substances and intendings and actions as events. The notion of psychological causation I will suggest attempts to avoid the picture of agents as substance-like. Rather, psychological causation has as its domain those changes and permanences in the flow of activity characteristic of living and, sometimes, self-conscious creatures. Such activity is organized by nature, by causal input together with internal structure and 'program', by conventions as adopted and internalized, and by conscious intentions and plans. Activity, rather than substances and states, should be the

category under which psychological causation is conceived. The flow, direction, redirection, and guided aspect of activity is the consequence of psychological causation; the activity itself, abstracted from its concrete exemplifications, is explained, if at all, by chemico-biological categories and laws.

The living, not unconscious, agent or the self-conscious, calculating agent, will be active. There will always be answers to "what is it doing?" When its activity is organized and directed, answers to this question require that psychological causation explain the action properties its activity satisfies. The issues here can be considered, first, from the perspective of the ontic causal processes and, second, from the perspective of explanation

Psychological causation is concerned solely the the psychological and actional properties and agent's circumstances, thinking and actions instantiate or satisfy. Its circumstances include the conventions in force, the historical and social setting, and the agent's linguistic practices. If its activity is fully intentional action, its psychological antecedents exhibit the conative properties present in the contents of its desires, wishes, preferences, felt obligations, etc., and the intentions these generate in particular circumstances. The processes of agency are channelled by and because of psychological causation into satisfaction of the properties expressed in its thinking. In a simple case of successful intentional action, the property expressed by "trying to fulfill a promise" will be satisfied by an undertaking, some physical movements and changes (e.g. handing over a check for $5) which itself satisfies the content of "intending to repay a loan of $5" as it occurs in the psychological processes of the agent. Here "satisfaction" is not merely property exemplification. In addition to satisfying the property of description "repaying a loan . . .", the physical undertaking or physical activity must be conceived or understood or thought of or represented *by the agent* as a means of or a way to satisfy this description. In successful, fully intentional actions (the only cases now being considered) psychological causation requires both the satisfaction of the description and the agent's belief or conception that his activity fulfills or satisfies the description present in his conative contents and cognitive states. (The semantic notion of satisfaction underlying these claims has been developed by R. Tuomela in *Human Action and Its Explanation,* 1977. See especially pp. 263–268.)

Psychological causation in the sense intended here requires, then, either property identity between the conation-cognition matrix, the undertaking, and the upshot if the action is fully successful, or property

coordination among these elements in less simple but successful cases of action. Thus, in the simpler cases, if the agent consciously intends to move the lever in front of him (expressed in a fully intention action sentence, say) and he successfully and fully intentionally does so, his undertaking exemplifies the property ". . . moving the lever" when this property was also part of his conative-cognitive complex in intending to move the lever. In more complex cases, if the agent intends to repay a $5 loan and successfully, fully intentionally does so, his undertaking exemplifies a property, "handing over a check", say, which is coordinate with a (typically means-end) property which was part of his conative-cognitive complex in intending to repay the loan. 'Basic action', as featured in some theories, requires here psychological causation involving property identity between intending and upshot. In 'non-basic action', psychological causation involves property coordination. Coordination of properties is based on our knowledge of the ways and means for tasks, the presence of skill and ability, and other forms of 'knowing how' and their internalization in routines of behavior. These, in turn, are expressed in the conventions and rules of behavior taught and learnt in social settings.

(There is an historical, partial parallel between the notion of psychological causation and some features of causation Descartes proposed as holding between an artisan and his works or between the Cartesian Creator and Sustainer and his works. The total efficient cause must exemplify the properties exemplified in its effects, either the self same property (e.g. heat) or another 'higher' property essentially coordinate with the defining property of the effect, e.g. the Creator does not exemplify the property of extension but does possess a 'higher' property, perhaps the property of fully comprehending the idea of extension, so that His works, the extended world featured by geometry, satisfy the geometrical properties coordinated with the higher property in the Creator. In Cartesian creationism, of course, the Creator produces not merely the possession of the essential properties of creation but the very existence of anything at all other than Himself. The Cartesian idea, nonetheless, is modeled on the relation between, say, sculptor and sculpture in classical thinking. It is not an accident that what I am claiming is a fundamental idea in agency and action should have a likeness to these ways of thinking. Psychological causation is meant to capture what seems apt in this classical and Cartesian picture.

Psychological causatioon is not only the dynamic and contentual connection present in deliberative intentional actions. It is manifest in contentual causal connections when agents routinely and habitually perform

actions which are not instances of essentially intentional doings. It is also present in purposive, primary intentional activities. At the self-con-

cates with the intentional connective (copula) which *are* intentions are 'preserved' in the descriptions under which the agent's actions are intentional. Thus, psychological causation differs from nomic causation as such, since in the last properties (or predicates in causal law antecedents) need not be reproduced or replicated in causal consequences. Lawfulness in natural causation takes up the slack in the absence of property identity or coordination. In psychological causation, property repetition and coordination takes up the slack in the absence of nomic necessity. Dynamic and productive causation itself can be present in both sorts of connections.

IV

The role of psychological causation in explanations of intentional action should be evident from the above. The scheme of intentional explanation at the level of common sense already suggests the presence of both property (content) duplication or property coordination and causation. Given that an agent could have done various things on a given occasion in that he enjoyed the freedom of indifference and had the ability and opportunity to do various things on the occasion, the questions can arise what was he doing? and why did he do that rather than some other thing? We can explain what he did by reference to his intention and the mode of its generation from his conative-cognitive condition. If we cannot see how a given intention could be the consequence of what were his conative-cognitive contents we may well find no suitable answer to what was he doing, though there will typically be some general answer in the form of a description of the physical or other changes attributable to him. But these will not be the result of psychological causation and will not cite such causation in their specification. We will attain a line on what the agent was doing, what *he was up to*, just in so far as we can see the psychological line or connection of causation which generates his intending. What he was doing *is* specified by his intention; his success depends in addition on other supporting conditions in the context in which he undertakes to fulfill his intention(s). And what he intends is generated in a variety of ways as we have seen. If detailed practical considerations were present, they will issue in an intention and will be *his* reasons or considerations just in case they psychologically generated his intending, for there may have been many considerations he entertained and others he could have considered, some of which would have been reasons for his undertaking, but which were not *his* reasons for acting, since they did not generate his intending. Though they were, suppose, reasons for his

undertaking, they were not his reasons because they did not yield, via psychological causation, his intending to undertake what he did. The class of possible justifying reasons is wider than the class of his reasons for adopting a given intention. We will have an account of why he undertook what he was doing in so far as we appreciate *his* reasons for such an undertaking, i.e., the considerations which actually generated his decisions and intendings.

Of course, not all intentional actions are preceded by practical reasoning in any of its forms. In the absence of appeal to any such reasoning there remain various modes of explaining what the agent was doing and why he did it. We saw in Chapter Eight that appeal to his circumstances and traits may be appropriate, in case we explain what he *must* have been doing by reference to the fact that he is the sort of agent who would do something of the sort. Clearly, we rely here on the notion of psychological causation and content transmission from what we take to be his traits, preference, and beliefs *to* his intendings and undertakings.

Further, when we find actions inexplicable, being unable to see how a given agent could have been doing what he appears to have done, or being unable to see how a given agent's apparent conative-cognitive matrix could yield an intention to do what he appears to have done, what we lack in such cases is an appropriate form of psychological causation. We are unable to understand how contents of the action and intention could have been coordinated with the apparent contents of the agent's conations; or, we are unable to see the property connections between the agent's conative-cognitive condition and the content we attribute to his undertakings.

Finally, in the absence of a causal component in explanation of intentional action and, in particular, in the absence of the feature of property replication and coordination, some cases in which an agent fails to undertake what he has a reason to undertake will be anomalous. If intentional actions are, roughly, just those the agent has a reason to undertake and thus undertakes to perform, then cases where reasons are evidently present but the agent fails to act (even where there is opportunity, etc.) will appear odd at best. The view which rests on the requirement of psychological causation in the generation of intendings and undertakings is not incapable of explaining such cases. The reasons and their contents were not sufficient to generate either the intending or the undertaking of an action whose description answers to the intention. No doubt this is not a full account of such failures. The point is, rather, that without the component of psychological causation no account of such failures seems possible.

The concept of agency and the implied notion of psychological causation has much the same role in the present view that willing or volition had in dualistic views. Willing was taken to preserve content. Dualism, of course, required that willing exhaust the agent's action, any bodily movement being a grace of nature. Rejecting this dualism, the present view takes undertakings (typically physical movements and, depending on the case, extra-bodily physical changes) as the physical aspect of agency. But there remains a role in the present view for something analogous to willing and volition. Willing, on this view, is the sophisticated psychological causation present in intentional actions. Fully intentional actions require, as remarked in Chapter Six, intentional activity, intentional further activity, and intentional objectivity. That is, full intentionality of action is present when the agent's will is fully engaged as to the activity expressed by the verb of action, the modification of the verb of action in further intentions, and by the intentionality of the object of the action. Willing, then, has the role already assigned to intending in the ostensive mode, when the content of the intention of such intending includes the full range of intentional activity, further intention, and intentional objectivity.

Furthermore, willing (or intending in the ostensive mode) is a psychological act and is 'episodic' or event-like and, thus, suitable as a cause. In short, it is a contentual cause. But few intentional actions are fully intentional in this special sense. An agent who intentionally raises his hand in greeting, say, may raise it 10 inches above his head but not have willed to raise his hand 10 inches above his head. He intentionally raised it, he intentionally raised it above his head, but while his willing produced the changes in bodily relations which is his hand's position 10 inches above his head, that specific position was no part of what he willed. His raising his hand above his head was fully intentional; his raising it exactly 10 inches above his head was not intentional. What an agent wills, then, gives a line of demarcation between intended action and accompanying consequences of action. Volition, in older theories, was designed as a concept for similar purposes. The action an agent performs is either the volition itself, as specified by its content, or the immediate upshot of the volition. All else is a consequence of action. On the present account, the action an agent performs fully intentionally is just those undertakings and physical changes which are marked by the content of his intention (or are changes whose properties are coordinate with ones marked by the content of his intention). Additional changes and events produced by the agent's undertaking are causal consequences of his intending and under-

taking; but since they are not generated by psychological causation, they are not intended consequences.

Willing is actualizing the capacity *for* fulfilling intentions. It is not actualizing the capacity to fulfill one's intentions, since what one intends is not necessarily achieved by one's willful undertakings. ("Exercising one's capacity to φ" carries the implication that one succeeds in φ-ing. "Exercising one's capacity for φ-ing" does not carry this implication. Exercising one's capacity for kicking goals does not imply kicking goals each time one tries. But a capacity for X-ing, say, does imply that one *can* succeed.) Willing, on the present view, is (1) bodily activity, (2) which exercises a capacity, and (3) which is subject to psychological causation. It is directed or controlled bodily will or willful bodily activity. An agent who cannot willfully move his limbs, who lacks the capacity to do so, lacks thereby the capacity for fulfilling his intention to move them. At best, such an incapacitated agent can form an intention to move his limbs, but in the absence of a capacity of the requisite sort, his desires, wishes and intentions do not connect via psychological causation with *his* moving his limbs. "Willing", of course, cannot be characterized apart from the notion of intention in one or more of its forms. In (2) above, "exercising one's capacity" is either a form of primary intentional action or a more complex, sophisticated undertaking. In both cases, it is intentional. What the notion of willing emphasizes is the psychological causal aspects of actions; hence, it has a place in the view of intentional activities and actions outlined here. On the continuum of intentional activity – primary, purposive, intentional activity *to* complex deliberative, deliberate actions – willing or volition characterizes only those self-conscious, deliberate, perhaps calculated cases where the requirement of contentual closeness between antecedents of action and action is at its highest.

V

Psychological causation is not only present in connection with intentional agency. It is present in perception as well, principally in perceptual judgment. The direction of causal connection is, of course, reversed in these two types of phenomena. In action, the content of thinking, intending, desiring, believing, and the conative-cognitives complex is replicated or coordinated in the content of action when successful. In perception, the content of physical states of the world (exemplified by light and color phenomena) is 'replicated' or 'coordinated' in the perceptual information expressed in perceptual judgments. This not the place to attempt to

express and defend a theory or perception and perceptual judgment. Rather, one important datum concerning perception should be noted; it will aid in the clarification of a similar feature of psychological causation in intentional action. The point of interest is that the vocabulary of perceptual judgment (especially color judgments) has a rather gross, underdetermined, approximate informational fit with the experiental content of perception. It is noteworthy and often noted that we do not have color-descriptive vocabulary for the actual determinate color content of experience. However fine grained our color vocabulary becomes with respect to shade, saturation and hue, intensity and any other such dimensions, the specific content of color experience is not expressed in the color words we use to describe and report such experience. We may call this the phenomenon of "information or content loss" in color judgments. We are taught color terms, as specific as you like, in connection with samples of a more or less standard sort, e.g. the color of the sky, of fire engines, mail boxes, flags, blood, grass in spring, etc. etc. What we actually encounter in concrete experience is some determinate, specific, demonstrative instance of red, blue, green, yellow, mauve, etc. And we have no words of the color vocabulary on a par with these color words for this perceptual content. At best, we may point and use "This shade of green" and similar demonstrative devices; there is no color word comparable to the standard color words in a language at a time for this content. We encounter *a green*, say, and report it with the term "green". (Of course, this is to say that color words are universals in the sense of classical thinking about these matters.) Yet the encountered, determinate color phenomenon is the referent of the appropriate terms in the color judgments these phenomenon (partially) cause, via psychological causation. There is, we can say, property replication or coordination with contentual loss.

This is also a feature of psychological causation in intentional action. Successful actions are some exemplification or other of the intentional complex, if they are the upshot of psychological causation. I have used the term 'coordination' of content loosely; as usual, these are matters of degree, matters of ineliminable vagueness *in the data*. Agents plan and intend to act; they direct their activities accordingly; the upshot is some action or other (a physical undertaking, say) which 'counts as' what they intended when their action is successful and intentional. There is 'contentual' or 'informational' loss in the best of cases, the most deliberate and calculated intentional undertakings where fit of content is crucial to whether or not agents performed as they meant. Put simply, the world as object of perceptual judgments is determinate and infinitely propertied;

but, perceptual judgment, in thought and speech, gains its descriptive facility by its generality and selectivity. The world as effected and changed by intentional agency is determinate and many properties, intentions, in adopting them and expressing them, are multiply realizable in events and changes some few of whose properties and relations must 'count as' fulfilling the content of intentions in case the events and changes are fully intentional actions. Planning and intending gain future directing facility by their generality, generality they share with future contingent statements as such. (Physical changes and events due to agency will also have determinate properties far beyond any envisioned, anticipated or planned by the agent and present in his conative-cognitive states. But this fact is not the focus of the present similarity between psychological causation in perception and action). Thus psychological causation makes room for the application of the intentional scheme of explanation in the various modes of its use we have noted; if psychological causation required a notion of property identity on all fours with strict, numerical identity, we should never be in a position to explain what an agent was doing (or did) by reference to what he would do, given our view of his preferences, personality and circumstances, for this mode of explanation is employed when we lack the advantaged position of detailed information of his operative conations and cognitions. Indeed, this is our typical position in relation to other agents. Hermeneutical understanding of intentional actions of others, thus, has a foundation in this style of explanation of intentional action (cf. *Human Action and Its Psychological Investigation* by A. Gauld and J. Shotter, 1977).

Finally, there is a significant difference between singular causal statements, (as Davidson explains them, for instance) and singular statements of psychological cuasation. The last are exactly the sorts of cases Davidson's analysis fails to capture. This is especially evident in light of the fact that psychological causation is so closely tied to the notion of intentional explanation. We thus must insist that specimen such as "The fact that he had been insulted and had not regained his composure and good humor by the time she returned with the wine he wanted caused him to argue with her selection of wines" contains "caused" which is best read as "causally explains the psychological reaction . . ." as Davidson suggests for similar cases (See Davidson's 'Causal Relations', *Journal of Philosophy* 64, 1967, especially pp. 702 f.).

VI

A minimal notion of "the agent" is required for the present view of agency. A Cartesian ego, transcendental self, and the like are replaced with a naturalistic conception of the agent as a physical locus of the causes which figure in agency and psychological causation. Here the agent is the changing locus of first-personal purposefulness in the flux of agency and activity; this locus is signaled, in English, in the use of the first person singular pronoun. As Wittgenstein urged, the pronoun is eliminable; some demonstrative, ostensive term or signal with the use of an indexical is, however, ineliminable. By means of the pronoun or alternative device with ostensive force, the locus of agency is indicated. (Compare the role of the first person singular pronoun in expressions of pain). Indeed, even this personal indicator has a limited role; in primary intentional action and much routine intentional activity, there is no strong first-personal focus. When first person reference is present in routine and habitual but intentional activity, it is often indicated *post actu*. In calculated, self-conscious, fully intentional actions, of course, the first-person of intention and agency is ineradicable. The semantics of simple and early developing intentional activity requires a locus of agency as subject of action attribution. The semantics of rationalized and calculated actions requires a person as the referent of attributions and as the self-referent of the intention. Strictly, then, it is intending (the attitude) in sophisticated cases of intending, further intending, and intentional action which requires a first-person subject of the attitude. Phenomenologically, a similar situation holds in connection with perceptual judging. Experience, characterized as perceptual believing, is much wider than self-conscious, articulate perceptual judgment. In the last but not the first, the perceiving subject is distinct and separable from the experiencing expressed in judgments. Experience and its causal fusion with the environment, its influence on reactive organisms, is a background against which explicit perceptual judgments are made. Activity and the origination of movement, active as well as reactive, is a background against which explicit intending is engaged. Yet a conception of the agent is required for articulate intentions in deliberative cases of intending.

A minimal agency view can best be set out by seeing how it deals with a pair of crucial issues. These are (1) how it avoids "all-out" agent causation, in a form of such causation which is claimed to be "foreign to science" (Davidson) and (2) how action is the action of the agent, *his* action, if not via the "all-out" causal agency of the agent's will? Consider these in turn.

(1) Brian O'Shaughnessy (*The Will*, Vol. 2, especially pp. 341–346) notes and explains admirably that what I have called "all-out" agent causation would be modeled on miracles. It would, thus, be foreign to more than science. Agent causation in the "all-out" sense is supposed to fill a gap between the agent's total antecedent conative-cognitive conditions, circumstances, ability, opportunity, etc., and his action. But what could this distance be, between the active, live agent's desires and reasoning and decisions and intendings, on the one side, and his undertakings on the other? Suppose an additional psychological-conative reality, act, is required to fill the 'gap'. Would there not be yet another bit of distance, now separating this conative act and the agent's undertaking? The "all-out" agent causation we have been discussing seems to be designed to full an unfillable gap, once we initially conceive of conative antecedents and undertakings as separable, on different sides of mind/body, inner/outer, self/world, etc. The "all-out" notion of agent causation appears fit only for a dualistic supposition melded to a full Cartesian notion of creative causation. On such a view, we would stand to our deeds as does the Cartesian God stand to the created world. With the rejection of this dualism or set of dualisms, the need for "all-out" agency is removed. We retain, of course, that aspect of Cartesian causation, distinctive of psychological causation, which requires property (content) duplication or coordination between causal antecedents and rational action.

(2) It is the rentention of psychological causation from the rejected "all-out" agency view which permits an answer to our second question; my action is mine in that it is the manifestation of my thinking, intending, and in general, my perspective and position in the environment. It is not the work of some part of me, my brain, my vital spirits, my mind, etc. It is not mine by virtue of being a part of something I own, as my sleeves are mine by virtue of being part of my coat, my echo is mine by virtue of its being a reflection of sound-waves of my voice, etc. To suppose that "all-out" agency is needed to make the natual upshot of my intending mine, is really to suppose that I stand to my intending and other conative-cognitive elements in an external relation; as if I must wait to see what my conations prove to produce; as if the absence of surprise were not an essential feature of intentional activity of the more or less sophisticated sort. The stress I have placed on physical capacities and learning (and consequent knowing how to do things) was a way of indicating the necessary tie of intending and undertaking. The possession of capacities and abilities, natural and learnt, rests on a reliable physical makeup and the

possession of the power of movement; intending, we have insisted, is a form of exercising such capacities. Manifesting or exercising physical (and some mental) capacities just is the undertaking consequent upon and controlled and directed by intending. The notion of "all-out" agency is designed to fill a necessarily vacant gap. The miraculous consequences such agency would require is a reductio of the underlying dualism between 'mental' conation and 'bodily' consequences. Living persons and animals who have not lost their native and acquired powers and skills are not models of such dualisms.

The "all-out" agency view can be seen as failing to recognize the way in which agency is constituted by conations. Necessarily, a living animal is disposed to act on its conations in accord with their contents. It is not so disposed in relation to conations of other animals or in relation to conations abstractly conceived. While we may share intentions, we cannot share our intendings, since we are two active lives. The "all-out" agency view models the 'relation' between the agent and its practical attitudes in rather the way one agent is related to the conations of another agent. But no undertakings would be the expression of my practical thinking and intending if they related to me in this external manner. And unless my practical thinking and intending was manifested in my rather than in another's undertaking, it would not be practical and conative at all. Living agents with remembered histories of preferences, decisions, commitments, and plans and with plausible futures in intentions, hopes, and plans are also active animals with powers and skills. As such, their conations, generic intentions and generated intendings here, now necessarily tend to manifest their psychology, personality, character (good or evil) in interaction with the world. We understand them as agents and comprehened their agency (what they do, would do, will do, etc.) via our appreciation of them as the physical arena of psychological causation. No more and *no less* is required for their agency.

The 'all-out' agency view is, at best, a deracinated version of the minority of cases of active agency in which we calculate, deliberate on alternatives, reject and adopt considered intentions self-conscious of our practical thinking. It is an uprooted version of self-conscious practical thinking in that it ignores the necessary physical and developmental background of agency; and it is a view built on a onesided selection of cases, in that it ignores the contentual causal relations manifest in purposive action, primary intentional actions, exercises of skill and ability and in routine, habitual or normal undertakings and their overwhelming success.

To recap and summarize: *Of course*, believing, intending, desiring, felt obligations, perspectives on forced alternatives, positive skills and powers are causally apt states, conditions, dispositions and episodes. *Of course*, beliefs, intentions, ideas of what is valuable or good for something, propositions concerning future contingencies, and rules and patterned descriptions of behavior are causally inert. But the last are among the kinds of structured feature the former have. As such, they are not explanatorily inept. (The quantitative values of forces are, as such, causally inert; the forces with those values are causally powerful. But the values are crucial for both explanation and, when otherwise possible, prediction as well. Properties are not fit for causation, though causally powerful states and events have their consequences in accord with the properties they possess.) Just as evidently, the causation of believing, desiring and intending is not Humean. Desiring to φ and thus intending to φ believing that one can do so by ψ-ing, is rationally connected with ψ-ing. Given that the agent has these features in amongst other relevant antecedents, it is less reasonable to think that he will not undertake to ψ than it is to suppose he might well ψ. The structured representational properties of the agent's psychological-conative conditions insure this result, i.e., that it is not just as reasonable to suppose he will not undertake to ψ as it is to suppose that he will. This is, if not Humean causation, Cartesian causation; it is a productive relation which obtains among formal (actual) things and events *and* among 'objective realities', conditions and episodes structured by content, by representations, as well. (Formal as well as eminent causation, in Descartes' words.) *If* physical causal laws of events range over unstructured event-particulars, psychological, explanatory statements range over structured, objectifying, contentual conditions. The last is no less causal for the differences. It is only that it is not Humean in the usual way as a consequence.

CHAPTER TEN

QUESTIONS AND SOME ANSWERS

> If action in general is pointless, then knowledge also is futile, and one belief is as good as another . . . Behavior which is beyond the subject's control is not action.
>
> C.I. Lewis

I

The view of intention and agency sketched in these chapters seems coherent in itself. It also seems to me to accord with the variety of data we have considered. While it is not as unified and systematic as some simpler theories, it attains a comprehensiveness which simpler views lack since simpler theories are often based on a narrower range of cases of action, for instance, on actions undertaken in fully intentionally pursuing conscious plans of action. As well, the present theory of agency is anchored in an activist, naturalist view. I have argued that the dualistic conception of the agent as an all powerful source of volitions is too fantastic to be taken seriously. Ethological ways of thinking should replace the Cartesian model of action. Evolutionary thinking should, together with human developmental studies, make a suitable view of action and agency sensitive to the continuity within all animal activity, from the reactive but controlled to the planned and calculated.

The activist framework has been elaborated in the account of psychological or contentful causation. Further, some of the sophisticated forms of conations (fully articulate intentions) and the types of explanation of intentional action have been sketched. I can now elaborate the view and better evaluate it by seeing how it can deal with a number of questions in the philosophy of action.

II

The questions which follow are in no particular order of importance, except that some of the answers require others.

These questions will not permit a straightforward answer in the stark

form in which they are put. Answers will depend on specifications as to the type of action, the type of intention (generic, ostensive, long-term plan, instrumental undertaking with further long term goal, etc.) and the scope of intentional objectivity in particular cases. So let us begin with some cases, specified in some of the indicated ways.

(1) Must an agent believe that he will do what he fully intends?

Suppose a wealthy grandfather is intent on leaving a large sum of money to each of a number of grandchildren, the balance of his estate to go to the preservation of the Alaskan Brown Bear population and its scientific study. He establishes a trust with this in mind. He performs a will with these specifics, etc. Of course, he does not know exactly how the trust will be managed; he does not know how it will preserve or increase his funds in light of the contingencies of national and world economy; and he does not know that the trust officers will be honest, etc., etc. There are many contingencies some of which could materialize and leave his grandchildren and the bears without support from his estate. So while he may well believe that no untoward contingencies will occur, he certainly is not in a position to know that they will be unforthcoming. Indeed, he might, in his paranoid senility, come to believe without grounds, that the trust is so mismanaged as to render his estate worthless, while he still harbors the desire and intention to provide for the interests in his will. Indeed, his paranoid belief might prove to be true. We can even imagine the case in which his groundless but true belief becomes a reasonable attitude, when the bank examiners and the trust department discover the embezzelment and so inform him. And this information need not generate desires and intentions contrary to the ones contained in his will. Nor need it alter his conative attitudes, for he may only now regret that his intentions are not likely to be fulfilled, but continue to hope for changes in his finances which will increase the chances of a suitable estate.

In this moderately complex scenario, where the fulfillment of intentions for the future depends on numerous future contingencies, there is at least one belief the agent of the will must have; he cannot have failed to believe that he undertook to provide for his heirs and for the bears in the drawing of a will. Negate this belief and the scenario is no longer coherent. Even in the case of an intention for the future whose fulfillment lies largely outside the conrol of the agent, the agent must believe that he will undertake to do his part in its fulfillment, if he fully intends it. "He fully intends φ but does not believe that he will do anything towards φ-ing" is incoherent. The first-person version is equally incoherent. Neither the

attribution nor expression of such a conative-cognitive condition admits of a consistent interpretation. Still, this does not exactly show that the agent must believe that he will φ, in case he fully intends to φ. The values of φ will range over cases which require the "friendliness of the circumstances" for their fulfillment. In those cases even competent and skilled agents may not believe that they will actually φ. But if they are able to φ they will naturally believe that they are likely to succeed when and if they undertake to φ. But if the values of φ are immediate exercises of physical powers, for instance, the power to do again what one has just done in the absence of any noticeable change in one's capacity (the power to type the next letter after the second *e* in "letter", say), then any suspicion that one will not do what one intends must be groundless. Of course, there can be conditions which remove even immediately realizable abilities, but their mere possibility (instantaneous paralysis due to onset of brain hemorrage or some such) will not count as a grounds for suspicion that one will not do what one is routinely undertaking in exercising a skill such as typing.

(2) Must an agent believe that he can do what he fully intends?

In the cases just described, if one believes that he will do what he intends and sets out to do, then he also believes that he can do what he sets out to do. Here the sense of "can" is the ability, power or skill notion. Much the same results hold in the external opportunity sense of "can". There is no reason to suppose that the opportunity to type the letter *r* after the last *e* in "letter" will be removed, say by sudden defect in the typewriter, since in has operated smoothly before now and nothing indicates a change. Nonetheless the opportunity might be lost by reason of a host of interferences. Where action involves the use of mechanisms which are subject to breakdown and where acting requires other supporting external conditions, and the absence of disabling conditions, e.g., one is not suddenly startled by a loud crash in the neighborhood, etc., the appropriate attitude is that one has no reason to think one cannot do as one sets out to do, though so-called objective certainly is not available. Fortunately, such extravangant certainty is everywhere unnecessary. It is suibable for omniscient agents. Fortunately for us we are not such agents. Typically, psychologically consistent agents will believe that they can do (and will do) what they fully intend. But there are two sorts of cases to be considered. In the cases of type (b) below, there is no room for non-bizzare doubt concerning ability and performance. Here are the case types:

(a) The agent intends to provide for her children's welfare by

means of investments and annuities, but she rather doubts that the return will be at all adequate, inflation being what it is, and so on. (A version of a case from Davidson).

(b) He intends to eat the mound of dessert in front of him. He couldn't harbor any doubts whether he is going to devour the load of pie and ice cream in front of him.

In (a) and its type the agent evidently can be in doubt whether she will do what she fully intends, for she can certainly doubt that conditions must be right for the long term future developments required for her success; she may well regret these prospects. In case (b) and its type neither an attributer nor the agent is in doubt and, it seems, the agent couldn't entertain a doubt – being on the verge of gluttonishly devouring his pie a la mode. Thus, the difference is that an agent can consistently intend to

that any conative condition is a case of wanting. Perhaps there is some such very general but uninformative notion of desire. But it fails to discriminate the fine grained differences which may contribute to a full understanding of the agent's thinking and acting. It seems advisable, then, to distinguish among the types of conation which satisfy the conative requirement in the scheme of explanation. Desire and wanting are not strictly required by the scheme; *some* conative condition is required.

(4) Must an agent know what he intends?

Since actions and actions directed on objects are intentional as conceived by the agent and since the agent's conception, picture, description, etc., can be defective in any number of ways, there is ample room for cognitive states short of knowledge. But since an agent cannot undertake to intentionally do a thing unaware that he is undertaking to do what he takes it that he is doing, there is no room for mistaken beliefs as to what one is intentionally undertaking. "I was unaware I was undertaking to . . ." and "I did not know that I was intending to . . ." are not open to coherent, non-bizzare, literal interpretations. In short, if an agent fully intends to do some thing, e.g., to φ x, then he must have a descriptive, referential, or other conscious awareness (i) that he intends something, e.g., to φ, (ii) that there is some thing, x, which he intends to φ, and (iii) that he intends to undertake something as the φ-ing of x. These are among the conditions for fully intentional, deliberate, conscious undertakings. As we have noted elsewhere such cases are not often exemplified in such stark clarity. For instance, there may be no actual thing to which an agent is referring (in thought of speech) when he has formulated the intention to find the man who just rounded the corner (it was an illusion or hallucination). Further, "to φ" might not be a standard action verb in the agent's native language and he may have a set of mistaken beliefs about what φ-ing is, how one φ's, etc. Thus, there can be mistakes in his conception of what he intends to do. Still, if an agent fully intends to φ, he must have *a* conception of φ-ing. None of these remarks go to the question of the coherence of the semi-technical idea of "unconscious intention", now in common thinking. For this, see (11) below.

(5) Is adopting an intention intentional?

Wittgenstein once remarked that "There might be a verb which meant: to formulate an intention in words or other signs, out loud or in one's thoughts. This word would not mean the same as our 'intend' . . . Yet another might mean: to brood over an intention; or to turn it over and

over in one's head." (*Zettel*, 49) To formulate, specify, think about, contemplate, brood over, compare with another intention, would each be activities. But adopting an intention seems at most one possible result of such activities. Adopting an intention, like deciding a matter, is an end-state, terminus, or conclusion of such activities as formulating an intention, contrasting it with other intentions, considering its pros and cons, etc. These activities are instances of essentially intentional mental (cognitive) activities. If they occur at all, if one engages in them at all,

etc., which accounts for the anomalous upshot. But I have insisted that this is an artefact of our commitment to the intentional scheme of explanation by means of which we try to understand and make reasonable to ourselves the thinking and acting of others. In some cases, we can find explanations, since such states and conditions as temporary moods, or underlying but previously unrealized traits or features of the agent's perception of his situation and so on, will be present and will account for the akratic consequence. But we have no sufficient reason to insist that all such cases of this type *must* be open to such a diagnosis. There simply are, it seems, situations in which an agent acts, say, dangerously or contrary to his perceived preferences or 'out of character' and so on. This sort of unpredictability, together with the variety of interacting elements in antecedent conations, cognitions, perspectives, moods and emotions, and traits of personality and character, make it implausible (and, I think, impossible) to find a formula or set of formulae which account for all cases of akrasia.

Of course, something like a formula for akratic actions is possible in certain kinds of cases. The problem is that we will be unable to apply formula-explanations except in rare situations, in which we possess almost complete understanding of the situation of others. One formula for cases in which intendings with appropriate content are generated by the appropriate antecedents but the intendings do not yield corresponding intentional action is just the view that there is or must be a physical, 'mechanical', breakdown in the causal background conditions in the agent's physical powers. Here the model of explanation for failures of action is rather like that for speech phenomena such as types of aphasia. We have more or less adequate physical explanations of these failures. Another formula explanation is the classical view of "passion dominating active reasoning". There do appear to be cases in which an agent fails to do what he/she fully intends because at the moment for action a desire or emotion (e.g., fear) of some sort interferes with the generation of the appropriate contentful undertaking. But not all cases seem to be of this sort; nor does the model of "being overcome by passion" appear generally appropriate, since we sometimes fail to act in accord with our best reasoning in the absence of anything as 'strong' or dominating as passions. Again, I conclude that no formula will explain every case we are likely to regard as akratic. For some cases, the best that seems available is a description of the situation, including the agent's conative, cognitive and emotional situation and the failure. Further, psychological causation is no more likely (less so?) to submit to an analysis in terms of necessary

and sufficient conditions than are other types of causal situations in the psychological and physical accounts of the breakdowns we call "akratic". In time the accounts of the last sort may come to include additional cases; this is an empirical issue.

(7) In the discussion of agency no attemted 'analysis' of power was provided. Does the present view admit of such a theory?

would have been realized in such movements. We will conclude that he has lost some power if in suitable 'test' circumstances he continually fails to move; but we may be wrong!

Again, we expect some explanation of the failure to F if we believe that the thing – artefact, natural system, organism, etc. – has an F power. We expect an explanation by reference to the thing, its circumstances, or a combination of the two. But our expectations are not necessarily satisfiable. In short, the causal power to F is just a feature that a thing has if *it* F's in suitable circumstances. If X, say, H's but is caused by Y in circumstances C to H, then X has a potential to H but it is not X's activity that is accountable for X's H-ing. This is not power; it is possibility or potential. The complication is, of course, that both X's power to F and its potenial of being H may depend on how X is structured. But X's powers are causally, internally dependent on X's constitution. In the case of agency, this is,we have argued, the natural evolutionary endowment and acquired abilities and skills of living agents in their historical and social settings. Abilities and skills, unlike other powers, are due to training, practice, trial and error and the rest of the notable conditioning situations which may obtain.

(8) How does the present view deal with the problem of causal inner waywardness?

We have introduced the causal and the psychological component in intentional action and reasoning to intentions in such a way as to deal with the issue of inner causal waywardness, in so far as it is open to solution on the empirical bases available. If the empirical theories (implausibly, I suspect) should succeed in uncovering a tighter causal account than the one endorsed by the scheme of explanation of intentional actions and activities, then parts of the problem of causal routes from cognitive-conative antecedents to intending and to undertaking will be solved. Empirical theories of the sort I have in mind just would be directed to this issue. Still, we can theorize on present bases and their projection. This we have done. The resulting account insists that both causal effectiveness of singular causal connections and psychological contentual duplication or coordination are required for the various degrees of intentionality our actions and activities exhibit. The issue of inner wayward causal routes seems to arise on the theories which over-emphasize the role of desire and under-emphasize the role of both contents of other conative conditions and the content of desire itself. I claim, then, that the proposed examples of waywardness (Davidson's climber example, Goldman's din-

ner guest case, etc.) are either not clearly cases of intentional actions or else are clearly cases of non-intentional actions. They are caused by the (or part of the) psychological antecedents but the mere causal effective-

assumed here that the intending is generated in a standard way in accord with the scheme of rational action; the agent had a reason for acting and acted because of it through an intention which his reasoning generated. However, his intending did not generate an undertaking which was a normal expression of his ability to do the thing intended or do what he thought necesssary and sufficient for the thing intended. The 'causal chain' from his intending took a course other than through his normal ability. Unknown to him, for instance, he had lost the normal ability but had, also unknown to him, been fitted with a prosthetic bit of neurological and muscular machinery, so that his intending activitated this mechanism and his purpose was realized. The facts, then, are these: He had reasons sufficient for φ-ing and intended to φ. In the circumstances, had he the normal ability, his intending to φ would have generated his undertaking to φ (or to ψ in order to φ, depending on other details of the case). He did not undertake to φ, since he lacked the appropriate ability. What did happen was that a prosthesis of some sort was activated by his intending; it produced his φ-ing. Now, he did not undertake to φ, intentionally or not. However, he *was intending* to φ. And, by grace of the fitted prosthesis, his intention to φ was realized. Where are we to assign the various forms of "intentional" to the elements of this case? Given these facts, it seems a matter of decision, so long as we keep the facts clearly in mind. I should say that (i) the agent was intending to φ, (ii) the agent did not intentionally φ, and, the crucial feature, (iii) the agent's intention to φ (expressed in (i)) was realized by the fortunate intervention of a causal route outside his abilities and their realization in *his* physical makeup. Note that if the agent comes to know of the prosthetic device he now has and comes to rely on this as on an ability, he will regard it as *his* in a way comparable to his attitude to his normal physical structure and ability.

Of course, there are imaginable cases in which the causal route does not go through a prosthesis but rather takes a route *wholly* outside the agent's physical structure, distant from a prosthesis. Suppose some necessary electrical potential and signal in parts of his brain are lost; the scientific community knows of this but the agent does not. So with considerable expense the community of engineers and scientists devise a method whereby the electrical impulse associated with his "having a reason to φ" is picked up by a device outside his head, magnified and transformed and sent to an orbiting NASA satellite where it is again transformed and sent to a receiver which in turn 'stimulated' the other end of the circuit in his head which is normally produced by circuitry

wholly in his head. This all happens rapidly! Maybe our agent only notices that he acts not quite as smoothly as when he usually undertakes to do what he intends. Now the facts are as follows: He intends to φ. He thinks he undertakes to φ because he intends to φ. He does undertake to φ, and he would not have undertaken to φ had he not intended to do so. But the route connecting his intending and undertaking is not the usual

the normal way up to the relevant time on Wednesday, when I pick up the vase and then drop it out of nervousness."

Now, was my releasing my grip intentional or not? (Peacocke's intuition is that it was intentional in relation to Monday's intention but not intentional in relation to Wednesday's.) Let us suppose the purported facts of the case are as follows: (1) I decide (on Monday) to break the vase on Wednesday. (2) On Monday it occurs to me that when I undertake to break the vase on Wednesday by picking it up, I will become very nervous and that my nervousness will likely cause me to drop the vase. (3) It aslo occurs to me on Monday that my nervousness on Wednesday and the likely consequences of this nervousness "will not prevent attainment of my goal". (4) It also occurs to me on Monday that I must stick to my intention to break the vase on Wednesday. I decide that I certainly shall break the vase on Wednesday. (5) I forget by Tuesday, say, the content in (3).

And (6) on Wednesday I pick up the vase and drop it out of nervousness. Now what is puzzling about the case and suggets its incoherence is the notion of the agent's thought that he must continue to intend (from Monday till Wednesday) to do what he has already decided on Monday. I have expressed this in (4), and I think (4) is unproblematic. But (4) might not capture the intent of the case. What could the thought that one must continue to intend what one has decided consist in other than continuing to think it something one wants to do? The idea behind the condition "I am successful in continuing to have the intention of breaking the vase in the normal way" seems to have no clear sense unless it means that some possible pros and cons concerning breaking the vase occur to our agent on Tuesday and Wednesday but none of the cons are sufficient for generating an intention not to break the vase. After all, intentions are not things agents can *manipulate* apart from adopting, rejecting, amending, clarifying, and making more definite in thinking and planning, etc. So if I have decided on Monday to break the vase and stick with this decision during Tuesday and Wednesday, then on Wednesday I do intend (as on Monday) to drop the vase and break it. The intention on Wednesday *is* the intention on Monday and Tuesday. Now, by (6), I drop the vase out of nervousness. If one takes this seriously, then dropping the vase was not intentional. Or, *the dropping* was not intentional. I was intending to drop the vase; what I intended thereby was to break the vase; the further intention, to break the vase, was fulfilled but not through my intending and undertaking in accord with the obvious capacity. Accordingly, this 'new' case seems only a version of the past cases. In reply to it the deci-

sions I prefer are (i) I decided (on Monday) to drop the vase and was intending to do so thereafter, (ii) nervousness intervened on Wednesday just when the generic intention had been reduced by the passage of time to the intention to drop it now, so the nervousness causally accounts for dropping it; my undertaking was precluded from its normal upshot. But, (iii) I got what I had intended, since the vase was broken. Perhaps, I didn't get the satisfaction of myself smashing Freud's vase. My intention was satisfied though perhaps I was not.

Thus, the last purported case seems only an elaborate version of the original case of 'inner wayward causes' proposed by Davidson. What

event featured correctly as an action, there will be an enormous number of additional true descriptions of it, each of which contains an action verb. And the multipliers can agree that some of the generated consequences of what they designate as the action will be action consequences, not merely additional description of the self-same action. Finally, the unifier theory in Goldman's version supplies no reason to object that there is no simgle action, in the unifier's sense of "action", i.e., in the sense of the underlying bodily agency attributable to the agent (in appropriate cases).

However, there are two central issues connected with the multiplier-unifier discussions which deserve further discussion. First, as Anscombe has insisted, the exprssion "x under the description d" is not a subject phrase (See her "Under a Description". Noûs **XIII**, 2, 1979.) Ontologically, we are not committed to two entities in the belief that action A under the description D is intentional and A under the description E is not intentional. Whatever action 'A' designates, say an undertaking in the form of a bodily activity of an agent, A is intentional under its description 'D' but not intentional under its description 'E'. Accordingly, "intentional under a description d" is a relational predicative device. Similarly, "tall in comparison to Theatetus" is a complex predicate which, say, Socrates satisfies, and "not tall in comparison to Plato" is one Socrates also satisfies. This does not generate contrary monadic properties; nor does it require two subjects such as "Socrates in comparison to Theatetus" and "Socrates in comparison to Plato", such that what one satisfies is "is tall" and the other its complement. Rather, the comparative and relational elements are predicated of Socrates. Similarly, as well, the relational feature expressed by "under a description" is predicative of an event, designated howsoever one designates it, and that event, if it is an action of an agent, can be both intentional under one description and not intentional under others.

The unifier view, then, takes account of this phenomenon. The self-same action can stand in contrary relations which contain, explicitly or implicitly, comparatives, contrastives, relations, and so on. Smith's act was generous in comparison to Jones's, miserly in comparison to Brown's; it was quickly done, in relation to Jones's, slowly by contrast to the Big Bang; it was more difficult than another undertaking of Smith's, less difficult than still others. Smith's action was intentional in light of its being his only hope to secure G, say, but not intentional in light of its being, say, an act which reduced his checking account by 15% exactly. If one insists that securing G and reducing the account are results or con-

sequences, not the action itself, then the same point can be made in connection with intentional and non-intentional consequences. The multiplier view, by contrast, individuates distinct acts via distinct descriptions, so it complicates matters when complexity is unnecessary.

Second, Anscombe has also rightly insisted that the question, But what is the action which has multiple descriptions? is a query from bare particular country. The idea behind such a question seems to be that there is (or must be?) some way of designating the action or event apart from its various descriptions and predicates; it is *the action*, thus indicated, which is *the* subject of the multitude of descriptions it satisfies. In some contexts, of course, there is something like this way of designation. If many suitable pragmatic conditions are satisfied, the action may be designated with indexicals or demonstratives. Even here it is a nice question as to the extent concepts and semantic relations are embedded in the conditions satisfying these pragmatic requirements. But this question

tion of the intention with which an action is undertaken. But motives contain a special belief or cognitive component, beyond the cognitive and know-how component of intentions. Revenge, for instance, is a typical motive. An agent with the motive of revenge believes that some offense has occurred, that certain parties are responsible for the offense and are inexcused, not exonorated, not justified for it, and that he is an appropriate agent of revenge. Apart from beliefs of roughly these sorts, his actions against the interests of another would not be acts of revenge, acts from the motive of revenge. Motives are in this way typically dependent on the environment in which they are held, including the beliefs and attitudes of the agent with a motive.

In addition to this belief component, an active motive, one yielding an intention and intending, will typically have either a strong desire component or another affective attitude such as the view that one *must* or *ought to* act. Motives, of course, need not be active. The presence of a motive is, for instance, only somewhat presumptive of guilt for an offense in law. It is the affective or desire component of a motive which gives it a role in the etiology of intentional action. Like the other elements in the intentional scheme of explanation, motives are cited in explanation of action. To appreciate that the agent acted from a motive of revenge is to put the right light on his undertaking and its consequences, in relation to his belief, attitudes or desires. Thus, by virtue of their cognitive, belief component motives are contentful and are fit elements in the psychological generation of intending. And by virtue of their affective component they often generate intending. How an agent will, if ever, undertake to act upon a motive, i.e., which means available for this purpose he selects, depends upon other beliefs as to appropriate means of, say, revenge. To understand the agent's actions as revengeful, then, is to understand their role as intentional means to his purpose, as he conceives the means and his purpose.

(11) The unconscious in the generation of undertakings.

Activities, styles of activities and actions arising from personality and character without conscious deliberation or deliberate intent appear to be what are often called unconscious intentions in popular, non-technical psychoanalytic discourse. However technically or untechnically one describes them, traits of character are such things as meanness, generosity, short temperedness, openness and sociality, secretiveness, and being fair minded, brutish, kind, a mixed class of tendencies, virtues and vices, etc. These terms characterize types of tendency learned or acquired in the

development of the personality, and characterize styles of thinking and responding. Activities and actions which realize these may be habitual or spontaneous. They are, however, purposive and may be purposeful as well. Realizations of them are subject to the intentional scheme of explanation; they are just what agents of that sort, with those traits and attitudes and beliefs, *would* do or think in those circumstances. Normally separable occurent desires, moods, and finely grained belief contents and so on are not required in explanation of actions which realize such traits. Morally relevant and social reactions of agents can be as habitual but as purposive as the exercise of physical skills and abilities. In these cases there is no need to construct elaborate contentual etiologies in applying the intentional scheme of explanation. Still, there is evidently a tremendously strong temptation to 'discover' elaborate thought processes and schemings in the unconscious life of the agent and to postulate events and processes of thinking, desiring, intending and motivation which are suitable causes of actions. There is no refuting these temptations!

Nonetheless, if there are simpler ways of understanding and explaining the cases which are supposed to be explained by unconscious motives, beliefs, and intentions and other unconscious schemings, these simpler accounts or models of explanation are preferable until there is either overwhelming evidence to the contrary or, perhaps, overwhelming conversion in society and culture to the contrary. Simpler forms of explanation are, of course, available in the views sketched in these studies. Unconscious event-like psychological acts are unnecessary in explanation by reference to powers, traits, abilities and circumstances. The notion of *psychological* explanation was introduced to augment the notion of mere causation, which does not require the same closeness of content among causal antecedents and consequents. Psychological causation caters to purposive, intentional activities and actions. We have outlined and defended the view of a continuum of cases, from the trait-like and circumstantial explanation of primary intentional purposive actions to the sophisticated, purposeful and deliberated fully intentional actions whose explanations require attribution (or avowal) of conscious deliberation or practical thinking. We have, then, available means for understanding some, at least, of the cases in which unconscious intentions seem needed in that we have (a) the scheme of intentional explanation, (b) psychological causal, in contrast to mere causal, connections and explanations, (c) and the notion of purposes in actions or the notion of intention in action, in contrast to the notion of explicit intentions with which acts are under-

taken. These, in the right mix, provide the resources for understanding at least some of the cases of 'unconsciously intended or motivated' activities. As well, of course, the sophisticated, cultural, psychological perspectives which now surround our thinking about actions, motivation and intention are the sources for the contents of 'unconscious intention'. My account of the last is meant to be independent of any such surroundings, however. Whether or not one thinks in Freudian terms, some intentions appear *in* the actions they help explain in accord with the special psychological vocabulary and thinking provided by psychological, literary conventions in force at given times in the history of culture. (cf. Ch. Eight, pp. 143–146)

More particularly, the notion of the intention *in* an agent's action is distinct from the notion of the intention with which the agent acts. By contrast with the first, the second can be said to be in the agent's thinking and planning and thereby in his action, under suitable conditions of psychological causation. An intention in an action is not typically in the agent's thinking prior to action. Rather, it is the content of the agent's ability or, in the present case of 'unconscious' intention, the content of personality, traits of character, or style of activity characteristic of the agent. We need not try to determine the locus of intentions in action, leaving it to theorists of psychology who might take up the suggestion I am putting forward as a model of so-called unconscious actional antecedents. For present purposes it is only necessary to note well the distinction between the types of intention here called " in action" and "intentions with which". Still more concretely, a person's action may exhibit the intention, say, to avoid controversy, but he is not acting with the intention to avoid controversy. He would not feature his own actions as avoiding controversy. He does not act with the intention of avoiding unpleasant encounters of controversy. It is only through critical self-evaluation and self-searching, for instance, that he comes to see this purpose in the pattern of his activities. Thus, the concept of an intention *in* action does not carry the implication that the agent can say or is aware of the content of his action. This is clear in both the developing cases of primary intentional action and the present target cases of agents whose sophisticated activities exhibit purposes which the agents do not have, in the sense of "intentions with which". If the notion of *intentions in actions and activities*, expressing or manifesting traits and propensities, rather than *intentions with which*, expressing sophisticated thinking and planning, is the notion of intention in 'unconscious intention', then some of the mystification surrounding this idea may be dispelled. (See for a few more de-

tails my discussion note, "On Unconscious Intentions", *Philosophy*, 48, 1973).

(12) How are nonintentional consequences or side effects of intentional undertakings explained so as to account for their lack of intentionality?

Obviously not everything an agent does in pursuing a further intention is itself intentional. In intending to signal a left turn he put his arm out the car window and struck a pine branch with his hand. So in intending to signal he struck a branch with his hand. But he put his arm out the window intentionally, while, suppose, he did not strike the pine branch with his hand intentionally. Though the agent did both, one was intentional, the other accidental, but both were done in intending to signal. How is this difference to be explained in the present theory?

It will not suffice to merely note that an agent might have thoughts such as "Q might happen," "Perhaps Q," etc. in the course of intentionally A-ing in order to R and, thereby producing both R and Q. The fact that some effects are merely pre-considered while some are intended, will not cover all cases in which we wish to distinguish accidental from intended consequences of our undertakings. Although substantially correct for a host of cases where R is the intended act and Q a considered possible side-effect, the condition is too strong for many cases of intentional action, e.g., spontaneous, unrehearsed activities as well as primary intentional actions, Further, the formula does nothing to explain the difference between doing A in order to . . .in contrast to, say, doing A while thinking that . . .Nor does it cater to side effects which, like the driver's striking the branch, are not thought of in advance of action at all.

This last weakness is easily removed. If the thinking, beliefs, considerations, and the like that figure in the agent's antecedent cognitive set have no coordinate or duplicate content in the effects of his undertakings, then their effects are not tied to his undertakings by psychological causation. Hence, on the present view, they are not intended outcomes of his undertakings. The important issue, then, concerns cases in which there is or appears to be a 'match' between the upshot of action and the cognitive-conative content in the antecedents of action. Here there does appear to be grounds for psychological causation and hence some grounds for intentionality in the upshot. This is the problem of considered but unintended consequences of intentional undertakings.

The notion of psychological causation again holds the key to a solution. Recall that psychological causation selects out the explanatory

properties in its causal antecedent and consequent from among the unlimited number of properties of the 'total antecedent state' of the agent. (The picture of the 'total state' is fishy, but let that pass; we need not be concerned with *effectiveness* for individuating such states at the present level of discussion.) Further, the causally relevant properties in an explanation of intentional action are those (or versions of those) which correlate the content of the cognitive-conative complex with the upshot of undertakings. Thus, in the abstract version of the present problem the property R but not the property Q is crucially exemplified in the conation of the agent. In the data, this fact is naturally expressed in the "in order to" relation (or another appropriate version of this relation). The property Q, while present in the host agent's thoughts and images, say, is not among the contents of his conations in the cognitive-conative complex. Hence, if Q is exemplified in the upshot of the agent's undertaking, it is not intentionally exemplified. If it is a mere causal consequence of his undertaking, we naturally call it a "side effect", and if its presence there is in some way untoward or unfortunate, it gets labeled "an accident" or, if useful or valuable in some dimension, it is "fortunate", etc.

(13) Reference in intentions.

As the contents of psychologically, causally productive and directing conations, intentions must have semantic features (particularly, reference).

But can an account of such representational states as intending be provided? I will say something on this difficult issue; more must wait on the background theory of mind in which these views on intention and agency will be set. As well, one can provide one's own account of the semantics of proper names and other designating instruments.

A naturalistic psychology of conative and other action relevant psychological conditions transparently does not require that the agent of such a psychology must enjoy even the rudiments of scientific omniscience in their representations of the world. While many psychologically active states – desires, practical believings, generic and ostensive intendings – contain reference to the agent's world, the nature of the *stuff* of the objects of these references may be (is?) radically unknown. Thus, agents such as our successful ancestors are capable of intentional activities in the absence of even the correct kind of concepts for a scientific and correct understanding of the way things are in the 'scientific image'. Intentions have as the domain of reference and description the actual but manifest world, the actual world as it puts in an appearance to agents with evolu-

tionarily designed perceptual powers and powers of agency. In the domain of 'common objects' there is reference and semantic success *and* failure. Too, there are strong existence presuppositions in conection with adopted further intentions (cf. Ch. Six, pp. 89–90). In *oratio recta* and indirect quotational attributions of further intention, special pragmatic rules are employed to block existence implications or presuppositions. Stress and emphasis as well as explicit negation of presupposition are required for the attributer of an intention with reference failure to express what the attributer takes the agent to have in mind to do. "Her further intention is to ship the newly arrived paintings to Mexico as soon as possible; that's why she is looking for regulations concerning art exports and shipping fees. But what she doesn't know is that her boss has destroyed all of the newly arrived works! (A seizure of utter madness, poor man!)" In expressions of further intentions of her own, the agent's ostensive references (in demonstrative intentions) are to be explained in whatever account of indexical, ostensive, or demonstrative reference one prefers, retaining the *de re* character of the resulting conative attitudes where needed to express the realistic example. Any such account will have a role for perceptual recognition and awareness of a direct but causally generated sort. No doubt much remains to be learnt and said about *recognition* in this context, since it admits of 'partial misfiring'. For instance, an agent recognizes the things, *a,* as a φ and "φa" is true, though the agent does not recognize *a* as a ψ, though "ψa" is also true. And this difference can be relevant to explanations of actions and forebearances. Her further intention in attending the meeting was to meet and speak with the Professor of Aesthetics who was also in attendance. Her further intention was to V *a*, under the description ψ, or in our schematic "She $V_\iota\psi_\iota$". But in the situation she recognized *a* as, say, φ (a visiting professor) but failed to identify *a* as ψ (Aesthetics Professor). If she, then, refrained from introducing herself to *a*, one explanation is that she did not recognize *a* as the one she had intended to meet. Or, suppose she did introduce herself to *a* but did not undertake to ask him about an idea of Malraux's on which he had written; one explanation is that she did not recognize *a* as the ψ, though she did recognize *a* as a φ, and hence she asked *a* if the Professor of Aesthetics was present.

Believings about and intendings about 'common objects', e.g., persons, their roles and salient features, physical objects, states and conditions of objects, etc., are typically read with both existential presuppositions and presuppositions of recognitional and referential success. In many realistic and actual circumstances we encounter, there is no recog-

nitional or referential failure, since these cognitive and conative successes do not require any degree of approach to scientific onmiscience on our part. Fortunately for our ancestors (and thus for us), no such references to 'scientific natures' and the like are typically required for successful believing, intending, desiring and acting. "It is simply the normal case, to be incapable of mistake about the designation of certain things *in der Muttersprache*." (Wittgenstein, *On Certainty*, 630). The very distinction between opaque and transparent reference and thinking occurs against this background, the same background against which we respond purposively, plan, decide and conduct complex activities of a more or less fully intentional sort.

Of course, there is a truth-permissibility condition on adopted intention attitudes of the following sort.

(a) "S intends to φa" or
(b) "S intends II the following: He (himself) to V (Ex) (φx at t)"

can be true when (i) there is no *a* which is φ, or (ii) nothing is φ at t. (V, here, represents a verb of action.) But a truth-permissibility condition does not require that (a) or (b) must be read opaquely or narrowly (i.e., without semantic properties of objective reference, etc.) in all tokenings of (a) and (b). Indeed, the views I have been pressing concerning generic and ostensive intendings require a principle of the following sort, reminiscent of Wittgenstein's attitudes in *On Certainty*:

(I*) Generic intendings contain intentions whose referring terms *necessarily typically successfully* refer
(1**) Ostensive mode intendings contain demonstrative terms with *necessarily typical* referential success.

In short, in the absence of these semantic features it would remain mysterious because miraculous that intendings should have the purposive and causal aptness they have in the actual world of actions and activities of agents. This is, at least, a 'transcendental argument' for these features.

(14) Free Action

Throughout I have made use of the notions of freedom of indifference and freedom of spontaneity. To repeat, an undertaking is free in the first way if the agent's psychological and developmental antecedents are such that the undertaking was one among some real options open to that agent on that occasion. An undertaking is free in the second way if it is the thing explained, in accord with the intentional scheme of explanation, by ineliminable reference to the so-called belief-desire matrix. Typically, one is

free in this way in case one does what one wants to do. Both of these components of free action are matters of degree. That is, an agent's freedom can be limited by diminished alternatives and it can be limited by conditions and circumstances which cause the agent to do something he does not prefer or something he is not happy to do. However, the mere presence of diminished alternatives or counter-preferenced undertakings does not of itself imply a lack of freedom. Before attempting to explain this last claim I need to introduce some clarifications concerning these notions of free action.

It should not be forgotten that the account of the belief-desire matrix and the explanatory scheme does not imply that an agent's intentional undertaking necessarily represents what the agent most wants to do. Desires can and often do play a secondary role in explanations which feature psychological causation; this is especially true, for example, where character traits and virtues play a leading role. Wanting is assigned this possibly secondary role on the present view because of the activist perspective in which the view has been developed. I have argued throughout that wanting appeared to be required as a special 'spring of action' because of the passivist presuppositions of traditional theories of action. Rejecting this presupposition, we were able to place wanting in what I take to be a more realistic and psychologically apt position among the conative antecedents of intentional actions.

Furthermore, it should be remembered that it is concrete undertakings which are to one degree or another free or unfree, and particular undertakings are judged free or unfree to one or another degree in light of both notions of free action. This is an important fact about judgments of freedom; it will help us to understand some features of our intuitions about limited freedom. For instance, if a person's undertaking was the only option avialable in the circumstances, then the action was forced and *in that way* less than fully free even if the agent acted from a desire and belief, etc. which generated an intending to do just that thing. Perhaps you had no psychologically real option but to spend the day performing a certain task; hence, you were not fully at liberty to do another thing. Still, you might well have wished to do just what you did; hence, you undertook the task without pain, happily, freely. Contrast this situation with one in which the taks was painful, hateful, done against one's wish and desire. Our intuition indicates a greater degree of unfreedom in this last case. And compare these cases with ones in which a person has psychologically real options, decides among them and happily does what he does. Here we judge the thing to have been done fully freely.

The idea of options or alternatives, is, of course, not just the idea of

so-called logically possible or even physically possible alternatives. Psychologically real alternatives are those which are (1) real alternatives in the circumstances and (2) function in the psychology of the agent in generating choices, decisions or intentions. Any number of complications in our judgments of free action are engendered by the fact that an agent can be misguided or mistaken about his/her real alternatives. The agent's subjective assessment and psychological-conative conditions might not square with the real alternatives in that the agent might take something to be an alternative when it is not or fail to take something to be an alternative when it is. Failure to have a realistic appreciation of one's situation, its alternatives, is, we think, a freedom curtailing condition. The causes of such failures is an empirical issue, though not necessarily a scientific, theoretical matter. Further, we have the idea that some persons are able to imagine and fantasize and wish and hope in greater scope and depth than others. There is a connection between freedom and the power or psychological capacity to imagine alternatives for oneself and integrate such imagined futures in one's practical thinking. Indeed, psychological health in the guise of ego strength is a freedom enhancing condition because, perhaps, it supports one's capacity to imagine options. The scheme of intentional action explanation and the psychological generation of intentions from the variety of action relevant antecedents caters to the role of imagination and 'mental' health. In turn, it caters to our intuition that a healthy agent, in this way, has greater freedom than, say, a neurotic.

It follows from this view of free action that the predicate ". . . is (was) freely done" is a relational predicate, requiring as arguments the concrete undertaking under psychologically apt accounting *and* a designation of the agent *qua* her/his psychological reality. There is, therefore, no such thing as a type of action called "free action". Further, since the attribution of relevant psychological realities (powers, proclivities, emotions, moods, beliefs and desires, traits, imaginings, and virtues and vices, to list a few) are defeasible matters, often requiring subtle assessments of what was undertaken, what the circumstances of action were, how the circumstances of action were, how the circumstances were viewed by the agent, and so on, the predicates "free" and "unfree" are matters of defeasible judgment as well.

CHAPTER ELEVEN

ACTIVITY AND DEVELOPMENT

> Animals manifestly enjoy excitement, and suffer from ennui . . . and many exhibit curiosity.
>
> Darwin

I

In this chapter I will survey just a sample of the work in developmental studies and ethological inquiries which give the activist view further empirical plausibility and a naturalistic setting. I cannot prove that it is the most (much less only) plausible conception of agency. I do not suppose such a claim can be proven. Still, given what we suppose we know about the structure and development of animals, including humans, this conception of agency seems plausible on every front. Of course, I also claim that the activist perspective has the empirical bases afforded by our usual understanding of agents and their deeds. In addition, it is set in the context of developmental and comparative ontogeny of behaviors. While our knowledge and theorizing are limited in these areas, the activist conception of action, the conceptions of power and ability, and the notion of psychological causation are designed to be understood in what seems the most plausible theories of behavioral development. In such theories, native proclivities, developmental stages of both continuous and noncontinuous sorts, periods of critical learning and developmental capacities, plasticity of behavior routines within the context of genetic makeup, and the crucial role of 'contingent' stimuli in development are all among the bits and pieces which supplement and support the present view of agency. Apart from the activist view, a host of data appears anomolous.

II

I have been at pains to emphasize a number of features of the activist view. Among these we should recall that in the domain of living animals there will always be answers to the question what is it, he, she doing now? The answers will feature, essentially, activity verbs at a molar level of

reference, beyond that of muscle movements, relative positions, and relative motions. In many cases the activity verbs will be purposive verbs. In some additional cases they will be verbs of purpose. The first are illustrated by "walking," "sitting," "arising," and so on. The second are instanced by "grasping," "calling to," "warning," "fleeing" and the like. Other such verbs name essentially intentional activities. E.g., their representative verbs include "looking up something in a text," "signing a legal document," and the like. So some verbs of purpose will represent contingently intentional and some essentially intentional activities. I have argued throughout that the primary intentional activities are on a continuum with sophisticated and calculated undertakings and performances.

What I have dubbed "psychological causation" or "contentual" or "contentful causation" comes into play in explanations throughout this continuum. At the level of primary intentional activities, merely purposive actions, it is employed in ethological and common sense attributions of purposiveness, means-ends connections, and in-order-to relations in characterizing the behavior of an animal or a developing human. Here the terms of psychological relations need not be (and mostly are not) available in the psychological processes of the target animal or human. But from the perspective of ethology as a science and from the point of view on attributions of developing behaviors and skills, these attributions are methodologically necessary tools. As well, in human development, articulate cognitive content and conative content (discrimination, recognition, acquired purposes, skills and abilities) typically follows rather than precedes primary intentional activity and rudimentary capacities and skills. As we develop into agents capable of selecting among and calculating concerning our powers of agency – what we now *can* do – psychological causation becomes the underlying phenomenon upon which rests our practical thinking *what* to do and *how* to do it. Because we now acquire alternative powers and become articulate as to what these capacities are, we can now consider what to do ourselves. At this point, it is appropriate to be more specific concerning the results of the developmental and ethological studies which support the ideas just expressed. (A representative source for these studies and results is K. Immelmann, *et al.*, 1981.)

The data and results of recent developmental studies are mixed and do not come close to settling a number of key issues, such as the 'nature/nurture' debate in some clarified sense, or the 'continuity/discontinuity of development' disputes. Still, the mixed data and plausible theories in

animal studies and developmental work do support the activist framework, and they appear inconsistent with the overly cognitivist, passivist perspective with which the activist picture contrasts. Nothing seems farther from the truth than Frued's alleged claim that the "mental apparatus endeavours to keep the quantity of excitation present in it as low as possible . . ." One need only consult the voluminous work on 'exploratory behavior' to become convinced that every animal (ourselves especially) seeks out new experience and varied environments by means of its own actions. The 'Nirvana principle' has little purchase outside elaborately sophisticated psychological contexts in particular culturally informed situations. Religious and psychological *views* about oneself, not oneself as agent, provide the context for such passivist ideas. (For a review of 'exploratory' behaviors, see D. E. Berlyne, 1960; C. G. Gross, 1968; and S. A. Barnett, 1975.)

We have reason to believe that a significant amount of learning and development is controlled by the native powers, the genetic character of members of the species. For instance, many behavioral capacities emerge relatively independently of both experience and age measured from birth. Developmental studies of term born and early birth animals and children suggests this in connection with vocalization, perceptual development, motor development and cognitive development. Thus the entire range of capacities seems to develop somewhat independently of early experience, conditioning and learning, etc. and appears to be a function of age measured from conception. Vocalization, for example emerges at the same stage, up to a point, in both early birth and term babies. In addition, the period development of vocalization is largely independent of perceptual capacities in that its onset is roughly the same in sighted and in deaf or blind infants. Motor development is similarly independent of these perceptual abilities. While sighted infants have obvious advantages, there is no more than a tendency towards earlier development; on the average they show rather slight advantage. In connection with motor development there is no evident significant difference between sighted and blind babies. In connection with walking alone there is only a slight advantage, on the average a month. But blind infants engage in hand regarding at the same state of development as the sighted. Sighted babies tested in darkness respond, as do blind infants, at the same stages with orientation to sound sources and grasping motions to sound sources. (See Andelson and Frainberg, 1974; Bower, 1977). Significant differences in development occur later, when visual information begins to be crucial in locating *kinds* of objects rather than merely locating

sources of sounds. And when vision is crucial to the presence of so-called contingent stimuli, i.e., stimuli defined in terms of voluntary control by the subject rather than in terms merely of the physical characteristics of the stimulus objects, the sighted will have decided advantages. But these are clearly behavior related advantages in that the sighted have the availability of perceptual (visual) feedback from the objects they manipulate and explore. The active interaction with stimulus objects is itself the crucial factor for contingent stimuli and their crucial contribution to learning and development. Behavioral deficiencies prove to be a greater disadvantage than blindness in this regard. These results surely fit the activist framework's conception of agency and power.

Recognition of the importance of contigent stimuli in development was evidently impeded by theoretical presuppositions of a behaviorist and passivist sort. If development and learning are primarily (or even exclusively) a function of stimuli, physically defined, then contigent stimuli, defined in relation to the manipulation and exploration by the subject, will only appear as a hinderance to controlled experimentation. And if stimuli are 'rigorously', physically defined, then the changes the infant or other subject makes in the stimulus condition should not make a difference in development and learning. The facts, evident now, are altogether to the contrary. Development and learning are enhanced and, in some cases, appear to be controlled by the capacity of the subject to move, alter and manipulate the stimulus condition and to actively alter its own perspective and relations to the stimulus object. Furthermore, the stimulus conditions defined in part by reference to the subject's interactions with its environment makes the physical characteristics of stimulus conditions much less significant, in some kinds of cases. The physical characteristics of the stimulus can vary greatly in every physical parameter; but if the infant controls the objects by his/her voluntary movements, learning and development progress. If the child or young animal is denied contigent stimuli, learning and development is slowed significantly or, even, in some cases, entirely blocked. (See R. Held and A. Hein, 1963; J. S. Watson and C. Ramey, 1972; J. S. Watson, 1981).

The central role of subject manipulation and control of stimuli in learning and development fits exactly with what one should expect from the perspective of the activist framework. It would be utterly remarkable if the natively active agent did not make use of its agency, power, control, and abilities in its development towards being a mature member of its species or group. Perhaps the most striking bits of data in connectin with exploratory activity and contingent stimuli concern their role in gener-

alization of neural function. For instance, there is some evidence that the availability of exploratory activity makes a difference to the recovery of vision in surgically traumatized animals. In a two stage removal of the visual cortex of rats, it was found that recovery of (partial) vision is possible for animals allowed to explore an environment between the first and second stages of visual cortex removal. Passive sightings of the same environment between the two stages of removal left rats blind at the conclusion of the procedure. (Cited by P. S. Churchland in her review of *Psychological Models and Neural Mechanisms* by Austen Clark, *Journal of Philosophy*, 1982.) (See also D. G. Stein, J. J. Rosen, N. Butters, eds., 1974).

We can note one final example where the role of contingent stimuli might prove to be important. Here its introduction into an experimental design can be easily seen. N. D. Henderson (*Journal of Comparative and Physiological Psychology* **72**, 1970, 505–511) studied mice of six inbred strains and hybrid type. Mice were selected for slow or for fast maze learning. Individuals from each genotype were reared in either a deprived environment (small cages, paucity of stimuli, restricted movements, etc.) or in an enriched environment (more nearly normal conditions). At six weeks all subjects were fasted and tested for food search abilities and maze learning. Among the results were that (i) mice from the enriched environment were quicker to solve the maze problem, independently of their genotype, and (ii) the genetically determined differences (slow or quick learners) were considerably more in evidence among all mice from enriched environments. However, Henderson evidently did not test subjects from the enhanced environment by separation into groups with and groups without active, exploratory (contingent) stimuli. One suspects that a repetition of this kind of study with the addition of this feature would reveal the fact that contingent stimuli availability is more significant than either genotype (for this trait) or mere enhanced, passive environmental stimuli. By hypothesis, it is not merely enriched physical stimuli conditions, but more importantly, enriched contingent stimuli conditions which account for learning and problem solving superiority, across genotypes. Henderson's results were preceded by similar work and like results on rats (D. O. Hebb), dogs (W. R. Thompson and W. Heron), primates (H. F. Harlo) and humans (J. Kagan). Any recent text in psychology can supply further details. (See, e.g. S. A. Barnett, *Modern Ethology*, Oxford U. Press, 1981).

Another feature of the activist framework is the rejection of the priority (in *any* reasonable ontogenetic, conceptual, or logical sense of 'prior-

ity') of the cognitive, theoretical, speculative over the active and the practical. By hypothesis, then, one should expect that development of sensorimotor and cognitive capacities should correlate; slow development in the first should be parallel to slow development in the last; accelerated development in the first should be found in connection with accelerated development in the second. In studies of the accelerated rate of sensorimotor development of infants in Uganda, it was found that the mean age of various features of language development were also highly accelerated in comparison to Western infants. The features of language development displaying cognitive acceleration were, in order of development, mere vocalization, name responding, use of a four word vocabulary, and development of comprehension of simple commands. (See M. D. S. Ainsworth 'Sensorimotor Development of Ganda Infants', in J. Oates (ed.,), 1979.) This is what one would expect on the view that active motor capacities and voluntary action developments are no less basic than so-called pure cognitive abilities. (See also M. Gerber, 'Developpement psychomoteur de l'enfant african', *Courrier 6*, 1965, pp. 17–19; M. Gerber and R. F. A. Dean, 'Gesell test on African children', *Pediatrics*, **6**, 1956, pp. 1055–65, cited in Oates, 1979). In short higher rates of acceleration of sensorimotor development in Ganda infants correlates with their accelerated rate of language development. It is reasonable to suppose that the conative, active capacities might well *contribute to* acceleration of language and cognitive development. Results of this sort motivate the rejection of the claim of priority of theoretical thinking over practical thinking.

Finally, in clasical and recent ethological theorizing the concept of a fixed action pattern has played an important role. A fixed action pattern (FAP) has the interesting feature that the stimulus is only a 'releaser' or 'trigger' of a complex patterned response which continues in an ordered way even in the absence of or the removal of the releasing stimulus. The Greylag goose will continue the standard pattern of behavior for retrieving an egg to the nest even when the stimulus object is absent, so long as it was present as a 'releaser'. The FAP continues to its 'fixed' termination in the absence of the triggering stimulus. (Cf. K. Lorenz and N. Tinbergen, *Zeitschrift für Tierpsychologie* **2**, 1–29, 1939). If the FAP of the nesting Greylag (and similar behaviors in other species) continues in the absence of the stimulus which causes its onset, perhaps standardized native and learnt patterns of human behavior (routines, skills, habits and the like) need not be supposed to require perceptual feedback mechanisms *as detailed causal stimuli* of such basic and primary intentional activi-

ties. The initial perceptual situation which is part of the cause of such activity need have no further role to play once an undertaking of this type is in progress. This would, also, explain the phenomenon we have emphasized as distinctive of intentional activities, *viz.*, *de re* knowledge without observation of the progress, stages, or termination and results of such actions. Basic motor patterns of human actions exhibit this trait; primary intentional action does so as well, once it is accompanied by learnt, articulate intention vocabulary on the part of developing agents.

III

Among the morals to be drawn from even a brief comparison of the activist perspective in the philosophy of action and the results of comparative, developmental, and ethological studies is that in philosophical theorizing about rational psychology and rational action we have too long emphasized the rationality of the rational animal, too long ignored or underestimated the role of the animality of rational animals. Indeed, it is on this basis that we have empirical and factual reasons for the claim of the equipriority (to put it paradoxically) of cognitive and conative functions, the interdependence of theoretical and practical thinking. We should let it strike us as remarkable that anyone should have thought that the products of the evolutionary past would be first and primarily cognizing, theoretical or speculative creatures. Jerome Kagan's theme in the following seems the better view when compared to the passivist and reactivist ideas characteristic of the recent theorizing,

> From Locke's *Essays on Understanding* to Skinner's *Beyond Freedom and Dignity*, we . . . have seen the perfectibility of man as vulnerable to the vicissitudes of the objects and people who block, praise, or push him. We have resisted giving the child any compass of his own . . . The maturation I have written about occurs because each physiological system or organ naturally exercises its primary function. The child explores the unfamiliar and attempts to match his ideas and action to some previously acquired representation because these are basic properties of the mind . . . (From Oates, 1979, p. 63)

BIBLIOGRAPHY TO CHAPTER XI

Adelson, and S. Frainberg: 1974, 'Gross Motor Development in Infants Blind from Birth', *Child Development* **45**, 114–126.

Ainsworth, M. D. S.: 'Sensorimotor Development of Ganda Infants', in Oates, ed., 23–34.

Barnett, S. A.: 1975, 'The Rat: A Study in Behavior' in *Modern Ethology: The Science of Animal Behavior.*

Berlyne, D. E.:1960, *Conflict, Arousal, and Curiosity.*

Bernstein, N. A.: 1967, *The Coordination and Regulation of Movements.*
Bower, T. G. R.: 1977, *A Primer of Infant Development.*
Churchland, P. S.: 1982, Review of A. Clark *Psychological Models Neural Mechanisms,* 1981, *The Journal of Philosopsy,* Feb. 1982, **LXXIX,** 2, 98–111.
Gerber, M.: 1956, 'Developpment psychomoteur de l'enfant African', *Courrier* **6,** 17–29.
Gerber, M. and R. F. A. Dean: 1956, 'Gesell Tests on African Children', *Pediatrics* **6,** 1055–65.
Gross, C. G.: 1968, In L. Weiskrantz (ed.) *Analysis of Behavioral Change.*
Held, R. and A. Hein: 1963, 'Movement-Produced Stimulation in the Development of Visually Guided Behavior', *Journal of Comparative and Physiological Psychology* **56,** 872–876.
Henderson, N. D.: 1970, 'Genetic Influences on the Behavior of Mice Can Be Obscured by Laboratory Rearing', *Journal of Comparative and Physiological Psychology* **72,** 505–511.
Immelman, K. G. W. Barlow, L. Petrinovich, M. Main (eds.): 1981, *Behavioral Development: The Bielefeld Interdisciplinary Project.*
Kagan, J.: 1979, 'The Importance of Simply Growing Older', in Oates (ed.) 577–63.
Lorenz, K. and N. Tinbergen: 1939, *Zeitschrift für Tierpsychologie,* **2,** 1–29.
Luria, A. R.: 1973, *The Working Brain.*
Schultz, D. P.: 1965, *Sensory Restriction Effects on Behavior.*
Stein, D. G. J. J. Rosen, N. Butters (eds.), 1974, *Plasticity and Recovery in the CNS.*
Watson, J. S.: 1981, 'Contingent Experience in Behavioral Development', in K. Immelman, *et al.* (eds.), 83–89.
Watson, J. S. and C. Ramey: 1972, 'Reactions to Response-Contingent Stimulation in Early Infancy', *Merrill-Palmer Quarterly* **18,** 219–227.

CHAPTER TWELVE

OVERVIEW

> The faculty which a being has of acting according to its ideas is LIFE.
>
> Kant

I

Native powers and their realizations form a continuum, ranging from controlled responses, habitual reactions, stylized and conventional routines, primary intentional actions *to* preconceived and articulated intentional actions, in their variety. A living, active agent will be doing some thing of various or these types at any point it its active life. I have been engaged throughout these chapters in articulating a theory around these claims. And I have claimed that one can fruitfully formulate and answer some of the questions in the philosophy of action from the resulting activist, naturalist perspective.

Included among our questions is the question of the nature of intentional action. Since the notion of intentional action has proved to be a general, cluster notion, no simple analysis is possible. Rather, I have had to examine cases of various types of actions and activities, the preconditions of each type, the nature and extent of the developmental, capability, psychological, and social antecedents of each type, and the variety of ways the intentional scheme of explanation caters to this collection of cases. At the base of the theory I have articulated, the notion of psychological or contentful causation is specifically suited to yield the control, direction, initiation of changes of activity and terminations of types of activity characteristic of the kinds of agents we suppose ourselves to be. The resulting theory differs from recent and current views of the nature of action.

As traditionally and currently rendered, intentional action is something like bodily movement with an appropriate cognitive antecedent. The agent, on this view, is merely the host of causal networks connecting appropriate antecedents with causally resultant bodily movements. The agent 'performs' an intentional action just in case the agent's desires and

beliefs, etc. cause certain movements to occur (in the right way). Strictly, the agent as such has no real role in the production of his or her performances.

A contrasting but less populated tradition places all the efficacy in the agent, the substantial individual, and proposes a contrasting notion of causation, typically called "agent causation". This view conceives of the agent as capable of producing movements, initiating changes in an otherwise passive substance, in accord with such causation. The beliefs and desires of the agent, on this view, are merely the forms and qualities which the agent is set on realizing in its movements.

In developing an alternative to these traditions I have been guided by both the insights and the failures I think both exhibit.

The most fundamental failure of these traditions is the lack of an adequate appreciation of the natively active nature of agents. There is a shared picture of agents as passive. Given this background picture, the problem for a theory of *action* will naturally include the problem of accounting for the absence of passivity, inactivity. The paradigmatic question, on these traditions, has the form of the simple question Why did agent A do X? where "X" ranges over actions. I have suggested throughout that the fundamental question has a less simple form, *viz.*, Why did A do X rather then W, Y, Z and so on? Accordingly, I have urged that we adopt an alternative picture of agents as active, so that the problem for a theory of action will naturally include the problem of accounting for the presence of a given type of activity in contrast to other types in which the agent could or would have been engaged. Thus, I have proposed that complex psychological antecedents are capable of supplying intentionalistic but causal accounts of the doings of active agents, although such antecedents are not suitable explananda of the movements of otherwise inactive things or substances.

In their accounts of intentional action, the dialectical interplay between the passivist traditions yields the result that *either* beliefs, desires and the like leave a gap between the agent and the action which seems unfilled by *any* added antecedents *or else* a unique and odd notion of causation (all-out agent causation) is required, with the problems which attend the introduction of a sui generis notion of this sort. As a way out of the dilemma, I have suggested that we (i) recognize with due theoretical accomodation the explanatory role of belief, desire and intention (and the remaining action relevant states or conditions), and (ii) also recognize the ineliminable activity of agents whose psychology and action we wish to explain and assess.

In light of (i) and (ii) I have proposed that intentionalistic accounts of agent's performances and forebearences employ a notion of causal explanation which is neither (in every respect) so-called event-causation nor the sui generis idea of agent causation. I have called this "psychological causation." The most primitive version of the idea of psychological causation is contained in such forms of assertion as "The agent did what *he* did *because* of *his* beliefs and desires and intention . . .etc." The contrasting primitive assertions, which fail to cater to (i) and (ii) alike, are "The action was performed because it is what the agent did" (the all-out agent case) and "These beliefs and desires, etc. caused this action" (the simple, event-causal case). In the last two forms of assertion, either the agent's psychology is not given an appropriate role or else the agent whose action it is, is not given an appropriate role. In short, I find a powerful clue in a paraphrase of the claim of T. H. Green cited as the front piece to Chapter Eight: It is not from the beliefs, desires, felt obligations, etc., but from the agent as affected by them, that the action proceeds.

The present view has origins in the data I have supposed any realistic account of agents and their actions must accomodate. Traditional views fare badly in light of the vast range of complex data. Suppose, for present purposes, that the traditional view is that intentional action is bodily movement initiated by appropriate cognitive antecedents. When placed in the context of a dualistic framework, this renders intentional action either miraculous or else an event restricted to one side of the inner/outer divide. I have tried to trace the main problems of this tradition to its exaggeration of the cognitive requirement for intentional action. The data, I claim, will not support such an exaggeration. Among the reasons, four have been specially important.

First, some cases of intentional action (purpo*sive* action, primary intentional action) are intentional but lack the typically requisite cognitive or 'mental' antecedents demanded on the traditional account. Second, the actual psychological antecedents of some moderately sophisticated but spontaneous and apt activities are not all that cognitive. In many actual cases, the requisite antecedents lie in the circumstances, the training, the development and the skills of the agent, including its perceptual and its motor capacities. Third, in fully, deliberated, intentional undertakings, the psychological antecedents require conative as well as cognitive conditions, *viz*, intendings, whose contents the agent acts upon. Fourth, in these sophisticated cases of fully intentional action, sophistication rests upon and develops from the less sophisticated capacities

and abilities of the agent as these are present in the agent's culture and life.

Nor is mental/non-mental dualism of the tradition bridged by the introduction of the notion of a super agent with *sui generis* powers of basic action. Such a notion does no justice to the complexity of types of action. It misconceives the connection between actions and their antecedents. And it abrogates the continuity among agents, from the non to the less to the more articulate and cognitively capable. The activist view, by contrast, proposes that agents are at the outset physically active, purposive initiators of changes in their activities and environments. They are engaged in and with the environment and sensitive to the 'information' it contains. At the very outset, they manipulate the environment to gather thereby 'information' which might be relevant to their needs and their curiosity. On this view agents are centers of vectors and directions of activities, from their native to their highly structured and preplanned performances.

In structured and fully intentional activities one might say that information is transmitted along the energic route of activity of the agent in accord with the information which *informs* the agent's psychology. If thereby the world comes to fit or match the agent's informed psychological states, intentional action has been successful. The notion of causation underlying this picture has roots in the idea that causes must 'resemble' their effects. This is, of course, a root only.

This theme is worth repeating once again. Agents are here conceived as causally powerful individuals, with contentful cognitive-conative states. The cognitive-conative contents of such states are not themselves causal agencies, since they are 'abstract' objects. Yet such contents are explanatorily crucial to the account of the agent's undertakings. Accordingly, the idea of psychological causation is first and foremost an explanatory notion. The vehicle of the crucial explanatory contents is, of course, the physical, energetic, active agent. In rational action, the contents of action relevant states are co-ordinated among themselves. In rational and realizable action, such contents are also co-ordinated with the environment in such a way that the world *can* come to conform to them. In successful rational action, the agent's psychological states generate co-ordinate or identical states with direct action potential. These are intendings. The content of intending (=intention) vectors the agent's energy and activity into undertakings. Fully successful intentional action is achieved when a friendly world cooperates in the production of an upshot which matches the directing contents of the agent's inten-

tions. The world is likely to be 'friendly' in case the rational agent has satisfactory beliefs and conforms his intentions to his beliefs. In another manner of speaking, 'information' is introduced into the agent-independent environment by the causal efficacy of the agent as that efficacy is formed by the informational content of the agent's psychology. The dynamics of these processes has been called, throughout, "psychological causation".

The notion of psychological causation I have elaborated is designed specifically to show how it is that agents are the initiators of changes, in and by virtue of their psychological antecedents and their abilities, without thereby being reduced to *victims* of those causal antecedents. In intentional action the cognizing and willful agent is not the mere, passive host of conditions which cause his intentional actions. Accordingly, the agent's belief-desire matrix is not the event-causal antecedent of his rational, intentional action. The matrix is, nonetheless, causally and explanatorily crucial to the very idea of intentional actions, since it is by virtue of the contentful causation of the matrix that the agent's living activities are sometimes intentional undertakings of his.

II

In their role as structured contents, intentions are distinct from intendings, though no agent *has* an intention which isn't an intending. Intendings, but not intentions, are causally apt. As we have noted, failure to note this fact accounts for the existence of some needless debates as to whether intentions (and reasons) can be causes. As 'abstract', formal features, intentions are not directly causally engaged. As real psychological conditions, intendings are vital, causally powerful, states and conditions. Thus, intentions provide the features, properties, forms which determine the kind and the extent of the causal, directive aptitude of intendings. It is intentions which go to explain what or how or why an agent is doing something, or was doing or has done some thing, when the agent's intending eventuates in undertakings and, then, successful performances.

Intentions have a variety of forms of expression or representation. Among these are "To do X with the intention to do Y," "I shall do X in order to do Y," "I am X-ing so as to Y," "I meant to X so as to Y," "I was Y-ing in X-ing," etc. Since intentions are logically apt, i.e. consistent or incompatible with other intentions and with beliefs, they have structures or forms. As we have noted, their from is not purely propositional (Cf.

Castañeda). First-person intentions require a practical connective, function or copula, which practically predicates an action or action-property of an agent, *viz.*, oneself. Since action properties themselves can be complex, with varying degrees of intention modification, an intention modifier (or predicate adverb, say) is also part of the structure of some, less simple, intentions. I have argued throughout that the psychological reality of agency requires that *intending* is not reducible to a combination of belief and desire; further, then, intending requires a content other than the content of cognitive states, and this is *intention*.

Underlying these views, only sketched in this "overview", are conceptions of the psychological reality of practical, living and vital states as well as conceptions of the formal and semantical notions of references and meaning associated with the contents of such states. Neither the philosophy of psychology nor the semantical theory have been in main focus in these chapters. Yet, there have been hints in the direction of some type of functionalist theory in psychology, especially in connection with practical states. There have been rather more than hints concerning the required semantical features of intentions. Large gaps remain. I do not intend to fill them here. However, it is useful to formulate, once more, some of the ideas which I suppose are required in order to fully defend the present theory.

For example, the conception of intentions (of the articulate sort required for fully intentional agency) requires self-reference. Self-reference occurs in unspoken thought and in thoughtful speech alike. Evidently, self-reference cannot be conceived along lines that might well be fruitful in the account of reference itself. So, for instance, self-reference cannot be conceived as a form of reference due to causal chains from antecedent reference, terminating in or extending back to prior causal rapport. Such a conception in the case of self-reference appears, at least, to generate problems of circularity and problems about the non-reflexitivy of causation. Intentional subjectivity (cf. Ch. Six) requires self-reference as a base mode of 'psychological causation'. The emergent thesis is, then, that in referring to oneself in the sense of ineliminably first-personal reference, the thought or expression which refers in this mode stands to oneself as the psychological effect on one's own thinking. Primitively, it is the agent him or herself who, informed by contentful states, thinks and sometimes avers the content expressed in forms such as "I shall . . ." This is just to say that in self-reference one thinks I-thoughts. There cannot be a further question as to what or to whom one is thus referring.

The present account of intentions also requires designation or *de re* reference in connection with the idea of 'intentional objectivity' (Cf. Ch. Six). The things upon which one intends to act are the objects of reference which are further identified as, say, *the F* or *an A* or *this S* which one intends to act upon. Normally, we accept relational readings or existential exportation, and the like, for attributions and expressions of intention. Additionally, the things by means of which one intends to act on something, the instruments of intended action as one might say, are the referents of parts of intention expressions. My hand, this finger, this tool, this instrument are objects of such intention parts. Thus, reference is presupposed by intention.

At the same time, we need to be reminded that in psychological-practical contexts of the sort introduced by intentional *subjectivity* and intentional *activity* and intentional *objectivity,* the typical de re/de dicto or relational/notional distinctions do not make a psychological difference, though they make a difference to the future success of one's projects in the world. Our theory tries, then, to accomodate the following facts: (1) We think about and propose to act in and upon the world with instrumentalities provided by our abilities and skills in the world. (2) We think, plan and intend in conformity with the way we think of and conceive of these items, i.e. under the descriptions of them present in our thinking. (3) Our thinking is assessed, and our practice engaged in accord with such thinking. (4) But, our success requires that the world itself suitably conforms to our ideas of it. Accordingly, we cannot be the sorts of agents we have supposed ourselves to be if we inhabit a pure notional world ('methodological solipsism'), and we cannot be the sorts of agents we suppose ourselves to be unless our thought and intention is genuinely *psychological*, that is,informed by the descriptions and notions under which we take the ambient world.

These ideas will not find theoretical accomodations in just any philosophy of psychology. More specifically, they will not be at home in a view which casts psychological states in the narrow, notional sense only. Nor will they be at home in a psychological theory which is merely behavioristic, externalist, or otherwise does no justice to intentionality and subjectivity. And, they will not find narrow functionalism a suitable cohort. Finally, they will not cohere with a meta-psychology which takes passive, cognitive states as the touchstone of psychology. These deeper issues lie beyond the present realistic account of intention and agency. To the extent that this account is correct, my results are data which any reasonable philosophy of psychology must accommodate.

INDEX OF NAMES

Ainsworth, M. D. S. 196
Aldrich, V. C. 2
Anscombe, E. 25ff, 58, 61–62, 77, 144, 175, 179
Aristotle, 13, 36

Barnett, S. A. 193, 195
Berlyne, D. E. 193
Binkley, R. 99
Bower, T. G. R. 193
Black, M. 99, 179
Butters, N. 195

Castañeda, H-N, 9, 115f, 127, 179, 204
Chisholm, R. M. 67
Churchland, P. S. 195
Clark, A. 195

Darwin, 191
Dean, R. F. A. 196
Davidson, D. 36ff, 99, 161, 162, 175, 178
Descartes, 46–47, 154, 155, 165
Dretske, F. 72

Feinberg, J. 99, 179
Frainberg, S. 193
Freud(ian) 133, 144, 188, 199

Galanter, E. 136
Gauld, A. 161
Gerber, M. 196
Goldman, A. 61–62, 179
Green, T. H. 122, 201
Gregory, R. L. 33
Gross, C. G. 193

Harlo, H. F. 195
Hebb, D. O. 195
Hein, R. 33, 194
Held, R. 33, 194
Henderson, N. D. 195
Heron, W. 195

Hume, 46–47, 165
Immelmann, K. 192

James, W. 20

Kagan, J. 195, 197
Kant(ian), 4, 6, 19, 46–47, 199
Kenny, A. 149

Lawrence, R. 57f
Lewis, C. I. 166
Locke, 53, 197
Lorenz, K. 196

Miller, G. 136

Oates, J. 196
O'Shaughnessy, B. 163

Panofsky, E. 58
Peacocke, C. 177f
Pribram, K. 136
Price, R. 84

Ramey, C. 194
Reid, T. 1
Rosen, N. 195
Ryle, G., xix, 107

Sartre, J-P., 60
Schopenhauer, 146
Shotter, J. 161
Stein, D. G. 195

Thompson, W. R. 195
Tinbergen, N. 196
Tuomela, R. 153

von Wright, G. H. 67
Watson, J. S. 194
Wittgenstein, 3, 7, 9, 15, 31, 69–70, 77, 89, 106, 110, 135, 162, 170, 187

INDEX OF SUBJECTS

Ability 5, 142, 149, 163
  'animal' 4
  and power 34, 93, 172–174
  and skill 150f
  and physical capacities 141, 146, 163, 175
  and 'can' 149
Action
  akratic 50f, 171f, 136f
  and action property 21
  and 'accordian effect' 93, 179
  antecedents of 124–26, 171f, 201
  'basic', 141
  and bodily movement 20, 136–138, 199
  and circumstances, 126f, 157, 201
  contingently intentional, 119, 139, 144–45
  and culture, 144–46
  free 188–190
  necessarily (essentially) intentional 29, 118, 139, 144
  and the passivist tradition 200
  and plasticity 151
  predicates 117–121
  repertoire 24, 132
  and side effects 184–85
  'springs of' 44, 46
  universals 114, 125
  'vocabulary' 23–24
  voluntary 23, 31, 93, 140
Activist perspective 18, 49–50, 141, 166, 191, 194, 197
Activity 29–30, 191–92
  exploratory 195
  intentional 89, 205
  organized 152f
  and passivity 21, 197
  and psychological causation 153f, 159
  the substrate of human action 2
Acts
  generic 10–11
  purposive 62

Agency 1, 158, 174, 191
  "all-out" 163, 200
  and ability 149, 152
  fully intentional 204
  and intention generation 152
  a model of 46f
  sophisticated 1, 24, 34–35, 146
Agent 148, 162–64, 199
  and causation 200f
  commonsense view of 6
Attitudes
  conative vs. cognitive 39, 47
  objects of 11
  practical 9,10
  propositional 9, 20
  pure intending 36f, 39

Beliefs
  and desires 133f
Bodily movement 136f

Causal
  efficacy of persons 1
  explanation of intentional action 16–17
  inner waywardness 174–179
Causation
  "foreign to science" 38, 47
  of agency 147
  and properties 155, 157
  psychological 150, 153–156, 160f, 165, 185, 192, 201f
  nomic 156
Commonsense 3–4

Desires 3, 34–35, 44, 49
  and "all-out" judgment 49
  as necessary and sufficient for intending 66f
  and 'springs of action' 46
Development 199f
  and comparative ontogeny of behaviour 191

and contingent stimuli 194f
and learning 22, 32
motor and cognitive 193
and natural capacities 23
and pre-deliverative actions 23
sensorimotor 196
and training 17–18, 32
Developmental (and comparative) Psychology 4, 22, 191f

Emotion 81
Ethology 15, 137, 140, 192, 196
Events
inner 74f
and intentional actions 15–16
under descriptions 180
Evolution 33, 47–49
Explanation 16, 123
and desire 34, 45–56
of idealized actions 124f
psychological (contentful) 155f, 175
rationalistic scheme of 126, 129, 146f
and social grouping 145
and traits 131–33

Fixed action pattern 196
Freedom 5
and alternatives 189
and imagination 189
of indifference 5, 34, 50, 188
of spontaneity 34, 188
Functionalism 205

'Hermeneutic circle' 129, 161

Intending 7, 20, 50
as dispositional 122, 144
generation of 127f, 156
paradox of 99f
and semantic aspects 123, 186–88
Intends
fully 167, 169
and belief 168f
and knows 170
Intention
in action 26, 188
the action universal view of 114
articulate 8,144
attributions of 57f
avowals of 70–71
collective 78,110
and *de re* reference 186f, 205
diffusiveness of 90f
expressions of 64, 69, 94–96, 102–03
for the future 40, 63–68, 98
further 33, 110, 127–28, 188
and inner events 74f
irrational 60, 133
mobilized 13, 102f
the modal view of 74–80
negation 100
non-personal 110
'objects of' 39
ostensive mode 13, 112, 120, 127, 188
as predictions 94, 96, 104–06
the propositional view of 111–114
standing 144–45
and self-reference 204–05
variable range of 118, 148
the variety of concepts of 3
with which 33, 78, 88–89
Intentional
action concept 18, 199
action and justification 22
action and side effects 184–85
copula 116f, 204
fully . . . action 13, 87–90, 201
objectivity 88–89, 113, 119, 205
predication 115f
subjectivity 205
Intentionally 84f

Judgment
agreement in 143–44
'all-out' 37, 44
prima facie 37, 40f

Linguistic practice 4

Motives
and affect 182
and belief 181
and explanation 182
'Multiplier' 16, 179f

Power 2, 168, 172–79, 199
and ability 149f
active 2
and control 151f
and development 191
to form intentions 151f
and natural capacities 32, 199
and role in explanation 34
and skill of 150f
Practical
knowledge 59f
reasoning 13, 21–22, 120
thinking 121, 152, 164, 197
and theoretical thinking 197
Primary intentional action(s) 18, 24, 28–30, 34, 59, 135, 155, 184, 192, 201
Purpose(s) 8, 34, 138, 140–41, 151

Trying 142–43

Unconscious 182–84
"Under a description" 16, 27, 85, 94, 121, 179, 180–81
Undertaking 141, 146, 182
Unpredictability (of action) 14–15, 135
'Unifier' 16, 179f

Volition(s) 13, 37–38, 127, 158
and willing 20, 53, 159