The philosophy of action

edited by Alfred R. Mele.

Published in this series

The Problem of Evil, edited by Marilyn McCord Adams and Robert Merrihew Adams
The Philosophy of Artificial Intelligence, edited by Margaret A. Boden
The Philosophy of Artificial Life, edited by Margaret A. Boden
Self-Knowledge, edited by Quassim Cassam
The Philosophy of Law, edited by R. M. Dworkin
Environmental Ethics, edited by Robert Elliot
Theories of Ethics, edited by Philippa Foot
The Philosophy of Mind, edited by Jonathan Glover
Scientific Revolutions, edited by Ian Hacking
The Philosophy of Mathematics, edited by W. D. Hart
Conditionals, edited by Frank Jackson
The Philosophy of Time, edited by Robin Le Poidevin and Murray MacBeath
The Philosophy of Action, edited by Alfred R. Mele
The Philosophy of Religion, edited by Basil Mitchell
Meaning and Reference, edited by A. W. Moore
A Priori Knowledge, edited by Paul K. Moser
The Philosophy of Science, edited by David Papineau
Political Philosophy, edited by Anthony Quinton
Explanation, edited by David-Hillel Ruben
The Philosophy of Social Explanation, edited by Alan Ryan
Propositions and Attitudes, edited by Nathan Salmon and Scott Soames
Consequentialism and its Critics, edited by Samuel Scheffler
Applied Ethics, edited by Peter Singer
Causation, edited by Ernest Sosa and Michael Tooley
Theories of Rights, edited by Jeremy Waldron
Free Will, edited by Gary Watson
Demonstratives, edited by Palle Yourgrau

Other volumes are in preparation

THE PHILOSOPHY OF ACTION

edited by

ALFRED R. MELE

OXFORD UNIVERSITY PRESS
1997

Oxford University Press, Great Clarendon Street, Oxford OX2 6DP
Oxford New York
Athens Auckland Bangkok Bogota Bombay
Buenos Aires Calcutta Cape Town Dar es Salaam
Delhi Florence Hong Kong Istanbul Karachi
Kuala Lumpur Madras Madrid Melbourne
Mexico City Nairobi Paris Singapore
Taipei Tokyo Toronto
and associated companies in
Berlin Ibadan

Oxford is a trade mark of Oxford University Press

Published in the United States by
Oxford University Press Inc., New York

© *Introduction and selection Oxford University Press 1997*

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press. Within the UK, exceptions are allowed in respect of any fair dealing for the purpose of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, or in the case of reprographic reproduction in accordance with the terms of the licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside these terms and in other countries should be sent to the Rights Department, Oxford University Press, at the address above

This book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
The philosophy of action / edited by Alfred R. Mele.
—(Oxford readings in philosophy)
Includes bibliographical references and index.
Contents: Actions, reasons, and causes / Donald Davidson—The problem of action / Harry G. Frankfurt—Trying (as the mental 'pineal gland') / Brian O'Shaughnessy—Acting for reasons / Robert Audi—Reasons explanation of action / Carl Ginet—A causal theory of intending / Wayne A. Davis—Practical reasoning / Gilbert Harman—Two faces of intention / Michael Bratman—Settled objectives and rational constraints / Hugh J. McCann—Intentional action / Alfred R. Mele and Paul K. Moser—Mechanism, purpose, and explanatory exclusion / Jaegwon Kim—Agency and causal explanation / Jennifer Hornsby.
1. Act (Philosophy) I. Mele, Alfred R., 1951- . II. Series.
B105.A35P49. 1997 128'.4–dc20 96–8682
ISBN 0–19–875174–5.—ISBN 0–19–875175–3 (pbk.)

ISBN 0–19–875174–5
ISBN 0–19–875175–3(Pbk)

1 3 5 7 9 10 8 6 4 2

Typeset by Best-set Typesetter Ltd., Hong Kong
Printed in Great Britain by
Biddles Ltd, Guildford and King's Lynn

CONTENTS

Introduction 1
Alfred R. Mele

1. Actions, Reasons, and Causes 27
Donald Davidson

2. The Problem of Action 42
Harry G. Frankfurt

3. Trying (as the Mental 'Pineal Gland') 53
Brian O'Shaughnessy

4. Acting for Reasons 75
Robert Audi

5. Reasons Explanation of Action: An Incompatibilist Account 106
Carl Ginet

6. A Causal Theory of Intending 131
Wayne A. Davis

7. Practical Reasoning 149
Gilbert Harman

8. Two Faces of Intention 178
Michael Bratman

9. Settled Objectives and Rational Constraints 204
Hugh J. McCann

10. Intentional Action 223
Alfred R. Mele and Paul K. Moser

11. Mechanism, Purpose, and Explanatory Exclusion 256
Jaegwon Kim

12. Agency and Causal Explanation 283
Jennifer Hornsby

Notes on the Contributors 308
Further Reading 309
Index of Names 313

INTRODUCTION

ALFRED R. MELE

A pair of questions are central to the philosophy of action. (1) What are actions? (2) How are actions to be explained? The questions call, respectively, for a theory of the nature of action and a theory of the explanation of actions. In short, philosophers of action want to know both what it is that explanations of actions explain and how actions are properly explained. One hopes that a full-blown philosophy of action will solve part of the mind–body problem and illuminate the issues of free will, moral responsibility, and practical rationality. But action is well worth understanding in its own right, as well.

Ordinary explanations of human actions draw upon a rich psychological vocabulary. In attempting to explain such actions, we appeal to agents' beliefs, desires, reasons (construed as psychological states by some), intentions, decisions, plans, and the like, and sometimes to deliberation or more modest forms of practical reasoning. Occasionally, we advert as well to finely distinguished traits of character and emotions. Traditionally, philosophers have refined and exploited this vocabulary in an effort to produce theories of the explanation of intentional human behaviour. An underlying presupposition is that common-sense explanations framed in these terms have enjoyed a considerable measure of success. We understand the behaviour of others and ourselves well enough to co-ordinate and sustain the wealth of complicated, co-operative activities integral to normal human life; and the understanding we have achieved is expressed largely in our common-sense psychological vocabulary. Even if the acceptability of this general approach to the explanation of action were taken for granted, however, we would face a variety of important questions about action and action-explanation. For example, inquiring minds still would want to know what actions are, what it is to act intentionally, what constitutes acting for a reason, whether proper explanations of actions are causal explanations or explanations of some other kind, and how the psychological or mental items (states, events, and processes) that are supposed to be explanatory of action are to be understood. These issues and the viability of the general approach just described constitute the primary focus of the essays reprinted here and of this introduction.

1. INDIVIDUATING ACTIONS

Question (1) 'What are actions?' suggests two others. How are actions different from non-actions? How are actions different from one another? I start with the latter, the question of *action-individuation*.

By the end of the 1970s, a lively debate over action-individuation had produced a collection of relatively precise alternatives: a fine-grained view, a coarse-grained view, and componential views. The first treats *A* and *B* as different actions if, in performing them, the agent exemplifies different act-properties. Thus, if I start my car by turning a key, my starting the car and my turning the key are two different actions, since the act-properties at issue are distinct. The second counts my turning the key and my starting the car as the same action under two different descriptions. Views of the third sort regard my starting the car as an action having various components, including my moving my hand, my turning the key, and the car's starting. Where proponents of the other two theories find, alternatively, a single action under different descriptions or a collection of related actions, advocates of the various componential views locate a 'larger' action having 'smaller' actions among its parts.

Interest in action-individuation has waned, owing partly to the development of a precise, detailed map of the conceptual terrain. Toward the end of an excellent chapter on the topic Carl Ginet remarks, 'the issue over the individuation of action, though sufficiently interesting in its own right, is not one on which much else depends'.[1] I am inclined to agree, and in this essay I proceed in a way that is neutral regarding the leading contending theories of individuation.

2. A CAUSAL APPROACH: OBJECTIONS, REPLIES, AND ALTERNATIVES

A popular approach to understanding both the nature of action and the explanation of actions emphasizes causation. Causal theories of action hold that an event's being an action depends upon how it was caused.

---

[1] Carl Ginet (1990), *On Action* (Cambridge: Cambridge University Press), 97.

These theories feature as causes such psychological or mental items as beliefs, desires, intentions, and related events (e.g. acquiring an intention to *A now*). If causal theories of action are on the right track, they provide a metaphysical underpinning for a popular causal view of the explanation of actions—the view that actions are to be explained, causally, partly in terms of items of the kind just mentioned. The conjunction of these two ideas—one about what actions are and the other about how actions are to be explained—may be termed *causalism*.[2]

Causalism typically is embraced as part of a naturalistic stand on agency according to which mental items that play causal/explanatory roles in action are in some way dependent upon or realized in physical states and events. A range of options are open in this connection: indeed, any viable solution to the mind–body problem that supports the idea that 'the mental' has a significant causal/explanatory role in action would, in principle, be welcomed by causalists. Causalism also is non-restrictive on the free-will issue. Although some causalists have endorsed compatibilism (i.e. the thesis that free will is compatible with determinism), compatibilism certainly is not entailed by causalism.[3] Provided that causation is not essentially deterministic, causalists can embrace libertarianism, the conjunction of incompatibilism, and the thesis that free will exists.[4] Some non-causalists are incompatibilists, but there is no entailment here either. Harry Frankfurt, a compatibilist,[5] rejects causalism in Chapter 2.

On an attractive causal theory of action, actions are like money in a noteworthy respect. The piece of paper with which I just purchased a Coke is a genuine US dollar bill partly in virtue of its having been produced (in the right way) by the US Treasury Department. A duplicate bill produced with plates and paper stolen from the Treasury Department is a counterfeit dollar bill, not a genuine one. Similarly, in typical causal theories, a certain event occurring at *t* is my raising my right hand at *t*—an action—partly in virtue of its having been produced 'in the right way' by certain mental items. An event someone else secretly produces by remote control—one including a visually indistinguishable rising of my right hand—is

---

[2] I borrow the term 'causalism' from G. Wilson (1989), *The Intentionality of Human Action* (Stanford, Calif.: Stanford University Press). (Wilson is a non-causalist.)

[3] Determinism may be succinctly defined as 'the thesis that there is at any instant exactly one physically possible future' (P. van Inwagen (1983), *An Essay on Free Will* (Oxford: Clarendon Press), 3). Some readers will wish to insert 'after the Big Bang' between 'instant' and 'exactly'.

[4] In A. Mele (1995), *Autonomous Agents* (New York: Oxford University Press) I develop three positions on free will, each of which relies on causalism: a libertarian position, a compatibilist position, and a position that is agnostic about compatibilism while advocating the existence of free will.

[5] Harry Frankfurt (1988), *The Importance of What We Care About* (Cambridge: Cambridge University Press), chs. 1–2.

试读结束：需要全本请在线购买： www.ertongbook.com

not a raising of my right hand by me, even if it feels to me as though I am raising my hand. Notice that the view under consideration does not identify actions with *non-actional* events caused in the right way.[6] That would be analogous to identifying genuine US dollar bills with pieces of printed paper that (1) are not genuine US dollar bills, and (2) are produced in the right way by the US Treasury Department. And, of course, so identifying genuine US dollar bills would be absurd.

The idea that actions are to be explained, causally, in terms of mental states or events is at least as old as Aristotle: 'the origin of action—its efficient, not its final cause—is choice, and that of choice is desire and reasoning with a view to an end' (*Nicomachean Ethics* 1139^{a}31–2). And it has enjoyed a considerable following. Owing partly to the influence of Wittgenstein and Ryle, this idea fell into philosophical disfavour for a time. The first major source of its revival was Donald Davidson's 'Actions, Reasons, and Causes' (Ch. 1 this volume).

There Davidson rebuts a collection of arguments against the causal approach. Some are still being advanced in various forms. I will take up two of these arguments and then turn to some further alleged problems for causalism.

(A) *The logical connection argument*

The most celebrated argument at the time hinged on the premiss that cause and effect must be 'logically distinct'. Because there is a logical or conceptual connection between an agent's wanting (intending, having a reason) to *A* and her *A*-ing, the latter cannot be an effect of the former; or so it was claimed.

Davidson's reply is incisive: causation is a relation between events, no matter how we describe them; the logical connections at issue are connections between event-*descriptions*. If *x*, the striking of the bell, caused *y*, the bell's tolling, our describing *x* as 'the cause of the bell's tolling' (as in, 'the cause of the bell's tolling caused the bell's tolling') plainly cannot change the fact that *x* caused *y*—the 'logical' connection between subject and predicate notwithstanding.

(B) *An argument from reasons-explanations*

The argument is straightforward. Causal explanations are lawlike; reasons-explanations are not; so reasons-explanations are not causal explanations,

---

[6] Cf. M. Brand (1984), *Intending and Acting* (Cambridge, Mass.: MIT Press), ch. 1.

and when we explain actions in terms of reasons, we are not explaining them in terms of causes.

Davidson agrees that *A* causes *B* only if 'some law covering the events at hand exists' (Ch. 1: 41). But he argues that the law need not be framed in terms of *A*-s and *B*-s; and he suggests that the causal transactions required for the production of action are lawlike, even though there are no (strict or suitably rigorous) psychophysical or psychological laws.[7] The idea evolves into 'anomalous monism' in later papers of his (1980, chs. 11–14; 1993[8]), a view characterized by the following three theses: (1) 'at least some mental events interact causally with physical events' (*Principle of Causal Interaction*); (2) 'when events are related as cause and effect, then there exists a closed and deterministic system of laws into which these events, when appropriately described, fit' (*Principle of the Nomological Character of Causality*); (3) 'there are no precise psychophysical laws' (*Anomalism of the Mental*).[9] The three principles jointly imply 'monism', Davidson argues: 'If psychological events are causally related to physical events, there must, by [2], be laws that cover them. By [3], the laws are not psychophysical, so they must be purely physical laws. This means that psychological events are describable, taken one by one, in physical terms, that is, they are physical events'.[10]

This raises some interesting and important questions. If as Davidson claims, the only (strict) laws are physical laws, why should we think that the *mental* features of physical events and states are causally relevant to the production of action? Are psychological or psychophysical laws required, after all, by typical causal theories of action or action-explanation? If there are no such laws, are we saddled with epiphenomenalism (the thesis that mental events and states are caused by, but do not cause, physical events and states) or worse?

Jaegwon Kim (Ch. 11) defends a principle—'the explanatory exclusion principle'—that, he contends, explains why neuroscience and 'vernacular psychology', including causal explanations of actions in terms of reasons, cannot peacefully coexist. The principle, in its metaphysical form, asserts that two distinct explanations of the same event 'can both be correct explanations only if either at least one of the two is incomplete or one is dependent on the other' (p. 275). He writes: 'Vernacular psychology and neuroscience each claim to provide explanations for the same domain of phenomena, and because of the failure of reduction in either direction, the purported explanations must be considered independent. Hence, by the exclusion principle, one of them has to go' (p. 281). Kim suggests that it is vernacular psychology that has to go, but only as a theory having the function of generating 'law-based causal explanations and predictions'—it may survive as a normative enterprise (n. 46).

---

[7] A psychophysical law links mental events or states (e.g. intentions) with physical events or states (e.g. bodily motions).

[8] Donald Davidson (1980), *Essays on Actions and Events* (Oxford: Clarendon Press); and id. (1993), 'Thinking Causes', in J. Heil and A. Mele (eds.), *Mental Causation* (Oxford: Clarendon Press).

[9] Principle 1 is quoted from Davidson (1980), *Essays on Actions and Events*, 208, as are the names of the principles. Principles 2 and 3 are quoted from p. 231. Incidentally, the intended root of 'anomalous' and 'anomalism' in the Davidsonian expressions reproduced here is the alpha privative prefixed to *nomos* in classical Greek, 'absence of law'.

[10] Davidson (1980), *Essays on Actions and Events*, 231.

Jennifer Hornsby (Ch. 12) advances a strikingly different view of things. She maintains that actions are events and that the explanation of action is causal explanation, and she argues that actions are not 'accessible' from the 'impersonal' (or neuroscientific) point of view. Actions, owing to their ineliminable connections with belief, desire, intention, and the like, and to the irreducibility that Kim mentions, disappear in a purely impersonal view of things. If there are actions, there are beliefs, desires, and intentions; and, Hornsby argues, the point of view from which actions are present is the only point of view from which they can be explained.

(C) *Arguments from causal deviance*

Instances of deviant causal chains raise difficulties for causal analyses of action, as Harry Frankfurt observes in Ch. 2. (This issue is not addressed in Chapter 1.) The alleged problem about *intentional* action in particular, as Paul Moser and I put it in Chapter 10, is that 'whatever psychological causes are deemed both necessary and sufficient for a resultant action's being intentional, cases can be described where, owing to a deviant causal connection between the favoured psychological antecedents and a pertinent resultant action, that action is not intentional' (cf. Frankfurt, Ch. 2).

The most common examples of deviance divide into two types. Cases of *primary deviance* raise a problem about a relatively direct connection between mental antecedents and resultant bodily motion. Cases of *secondary deviance* focus on behavioural consequences of intentional actions and on the connection between these actions and their consequences. The following are, respectively, representative instances of the two types of case:

> A climber might want to rid himself of the weight and danger of holding another man on a rope, and he might know that by loosening his hold on the rope he could rid himself of the weight and danger. This belief and want might so unnerve him as to cause him to loosen his hold [unintentionally].[11]
>
> A man may try to kill someone by shooting at him. Suppose the killer misses his victim by a mile, but the shot stampedes a herd of wild pigs that trample the intended victim to death.[12]

Instructive attempts to resolve the problems such cases pose highlight four points (see Audi, Ch. 4; Brand 1984[13]; Harman, Ch. 7; Mele and Moser, Ch. 10; Searle 1983[14]; Thalberg 1984[15]). (1) A necessary condition of an *A*-ing's being an intentional action is that it be an action, and in many cases of deviance the pertinent event seems not to be an action. For example, the climber's 'loosening his hold' is more aptly described as the rope's slipping from his trembling fingers. (2) An analysis of action may preclude there being a gap between the agent's action and its pertinent psychological causes, thus eliminating the possibility of primary deviance. If, for example, every intentional action has the acquisition of a 'proximal' intention (an intention to do something *straightaway*) as a *proximate* cause, there is no room between cause and action for primary deviance.[16] (3) Intention (or one's preferred psychological item) should be given a *guiding* function in the development of intentional action. (4) An action's being intentional depends upon its fitting the agent's conception or representation of the manner in which it will be performed—a condition violated in the shooting case, standardly interpreted. For development of these ideas, see Chapters 4, 7, and 10.

Some attempts to come to grips with causal deviance invoke the thesis that all intentions are self-referring. Gilbert Harman puts the view succinctly: 'the intention to do *A* is the intention that, because of that very intention, one will do *A*' (Ch. 7). The basic idea is that intentions specify how they will lead to action, and that if they do not lead to action in (roughly) that way—as in many cases of deviance—they have not issued in intentional action. I have argued in response that incorporating non-intention-referring plans into intentions is no less successful in blocking deviance and avoids problematic commitments of the thesis that every intention refers to itself (Mele 1992*a*[17]). For a critical reply, see Harman 1993.[18]

---

[11] Ibid. 79. [12] Ibid. 78.

[13] Brand (1984), *Intending and Acting*, ch. 1.

[14] J. Searle (1983), *Intentionality* (Cambridge: Cambridge University Press).

[15] I. Thalberg (1984), 'Do Our Intentions Cause Our Intentional Actions', *American Philosophical Quarterly* 21: 249–60.

[16] '*Proximate* cause' may be defined as follows: *x* is a proximate cause of *y* if and only if *x* is a cause of *y* and there is nothing *z* such that *x* is a cause of *z* and *z* is a cause of *y*.

[17] A. Mele (1992*a*), *Springs of Action* (New York: Oxford University Press), ch. 11.

[18] G. Harman (1993), 'Desired Desires', in R. Frey and C. Morris (eds.), *Value, Welfare, and Morality* (Cambridge: Cambridge University Press).

A third variety of deviance ('tertiary deviance', Mele 1987*b*[19]) merits attention. It was originally brought to bear against Myles Brand's claim that '*S*'s *A*-ing during *t* is an intentional action [if] (i) *S*'s *A*-ing during *t* is an action; and (ii) . . . *S* has an action plan *P* to *A* during *t* such that his *A*-ing is included in *P* and he follows *P* in *A*-ing'.[20] Here is an example:

> Fred is taking a machine-readable multiple choice test. His strategy is to circle on the question-sheet the identifying letters next to the answers that he feels certain are correct and then, after all such circling is completed, to fill in the corresponding spaces on his answer-sheet. . . .
>
> An hour has elapsed, and Fred is reading the forty-fifth question. He is confident that the answer is 'bee', which word appears next to the letter '*a*' on his question sheet. However, as a result of an understandable momentary confusion, he circles the letter '*b*'. As luck would have it, '*b*' is the correct answer. Later, when filling in the answer sheet, Fred looks at the circled '*b*' under question 45 and fills in the space under '*b*' on his answer-sheet—intending thereby to provide the right answer.[21]

Fred provides the correct answer to question 45 by filling in the space under '*b*' on the answer-sheet. But his providing the correct answer, in light of the deviant history of his intention, seems too accidental to be intentional. Still, Fred did what he intended *at the time* to do then—namely, to provide the correct answer by filling in the space under '*b*' on the answer-sheet. Further, his so doing was initiated and guided by his intention to do so, an intention incorporating a plan that he followed. (Notice that the causal deviance occurs *prior* to Fred's forming the pertinent intention and therefore is neither primary nor secondary deviance.) In Chapter 10, Paul Moser and I develop a causal analysis of intentional action designed to be immune to all three sorts of deviance.

(D) *Frankfurt's argument*

In Chapter 2, Harry Frankfurt advances an argument against causalism from what might be termed 'passive action'. He writes:

> A driver whose automobile is coasting downhill in virtue of gravitational forces alone may be entirely satisfied with its speed and direction, and so he may never intervene to adjust its movement in any way. This would not show that the movement of the automobile did not occur under his guidance. What counts is that he was prepared to intervene if necessary, and that he was in a position to do so more or less effectively. Similarly, the causal mechanisms which stand ready to affect the course of a bodily movement may never have occasion to do so; for no negative feedback of the sort that would trigger their compensatory activity may occur. The behaviour is purposive not because it results from causes of a certain kind, but because it would be affected by certain causes if the accomplishment of its course were to be jeopardized. (Ch. 2: 48)

---

[19] A. Mele (1987*b*), 'Intentional Action and Wayward Causal Chains: The Problem of Tertiary Waywardness', *Philosophical Studies* 51: 55–60.

[20] Brand (1984), *Intending and Acting*, 28.

[21] Mele (1987*b*), 'Intentional Action and Wayward Causal Chains', 56.

In the absence of a desire or intention regarding 'the movement of the automobile', there would be no basis for the driver's being 'satisfied' with the speed and direction of his car. So we may safely attribute a pertinent desire or intention to the driver, whom I shall call Al. What stands in the way of our holding that Al's acquiring a desire or intention to coast down hill is a cause of his action of coasting, and that some such cause is required for the purposiveness of the 'coasting'? Even if Al passed out momentarily at the wheel and then, upon regaining consciousness, noticed that his car was moving smoothly down hill, his allowing this to continue to happen, owing to his *satisfaction* with the car's speed and direction, depends (conceptually) on his having some relevant desire or intention regarding the car's motion; and prior to his allowing the continuation he is not purposively or intentionally coasting down hill—he is merely being carried down hill. We are left with the same question.

Perhaps, unbeknownst to Al, the brakes, accelerator pedal, and steering-wheel are no longer working: his car would continue moving as it is even if he were to lack the desire or intention in question. But then Al is not performing an action of coasting down hill—he is merely being carried along by a vehicle over which he has no control. And if he is performing no such action, he is not purposively or intentionally coasting, even if he thinks he is. (Notice that the claim that Al *is* purposively coasting in this case is at odds with Frankfurt's own position on purposive behaviour; for it is false that Al would have corrected the car's course if he had deemed it unsatisfactory.)

So suppose that the car is in normal working order and that Al knows how to operate it, is not paralysed, and so on. Then it is natural to say that Al is coasting in his car (or allowing the car to continue to coast, in the scenario in which he wakes up in a moving car) *because* he wants to, or intends to, or has decided to—for an identifiable reason (e.g. to conserve petrol). And the 'because' here is naturally given a causal interpretation. In a normal case, if Al had not desired, or intended, or decided to coast, he would not have coasted; and it is no accident that, desiring, or intending, or deciding to coast, he coasts. So, setting aside general worries about mental causation, it looks as though Al's coasting does have a mental cause.[22]

---

[22] Consider a scenario that differs from the preceding one only in the following respect and in ways entailed by the change. If Al had not decided to coast, he would have been utterly indifferent about the motion of his car—in which case he would have done nothing to alter the

Frankfurt might reply that even if Al's coasting has a suitable mental cause, his coasting is purposive 'not because it results from causes of a certain kind, but because it would be affected by certain causes if the accomplishment of its course were to be jeopardized'. The idea is that what accounts for the purposiveness of the coasting is not any feature of how it is caused but rather that Al 'was prepared to intervene if necessary, and that he was in a position to do so more or less effectively'.

This reply is problematic. Imagine that, throughout the episode, Al was satisfied with how things went and did not intervene. He decided to coast and the coasting was purposive. Imagine further that although Al intended to intervene if necessary, an irresistible mind-reading demon would not have allowed him to intervene. If Al had abandoned his intention to coast or had decided to intervene, the demon would have paralysed Al until his car ran its course. The coasting is purposive even though Al was *not* 'in a position to [intervene] more or less effectively'. And this suggests that what accounts for the purposiveness of Al's coasting in the original case does not include his being in a position to intervene effectively. There are, moreover, versions of the case in which Al's coasting is purposive even though he is not prepared to intervene. Suppose Al is a reckless fellow and he decides that, no matter what happens, he will continue coasting. He has no conditional intention to intervene. Even then, other things being equal, his coasting is intentional and purposive.

Regarding causal theories of action, Frankfurt writes: 'it is beyond their scope to stipulate that a person must be in some particular relation to the movements of his body *during* the period of time in which he is presumed to be performing an action. The only conditions they insist upon as distinctively constitutive of action may cease to obtain, for all the causal accounts demand, at precisely the moment when the agent commences to act' (Ch. 2, 43). However, some causalists have argued that the causal role of the mental items on which they focus (intentions, or beliefs and desires) includes a sustaining and guiding function (see, e.g. Audi, Ch. 4; Brand 1984; Mele 1992*a*; Mele and Moser, Ch. 10). And they may hold that Al performs the action of coasting down hill in his car partly in virtue of his car's motion's being causally sustained by an intention or desire of his to coast.[23]

---

car's course and the car would have continued coasting. In this scenario, it is false that if Al had not decided to coast, the car would not have continued coasting. Even so, Al's deciding to coast is plausibly regarded as a cause of the continued coasting. Compare: *X* dialled *Y*'s phone number at *t*, but if *X* had not done so, *Z* would have done so (at *t*). *X*'s dialling is a cause of *Y*'s phone's ringing at *t1*, even though the phone would have rung at *t1* if *X*'s dialling had not occurred.

[23] Is this suggestion threatened by the scenario that features a non-interfering demon? In that case, Al's car would have continued to coast even if Al had not continued to intend to coast.

(The other sort of behaviour to which Frankfurt alludes in the quoted passage is subject to a similar causalist treatment.)

(E) *Ginet's alternative to causalism*

Non-causalists tend to agree with causalists that actions are to be explained in terms of reasons, intentions, and the like. But they disagree with causalists about the nature of the explanations offered. For non-causalists, desires and intentions may be loci of agents' purposes in doing what they do intentionally, but they make no causal contribution to intentional action. Obviously, some non-causalists are moved by considerations of the sort discussed in subsections *A–D*. Some are moved as well, or instead, by considerations of free will and self-knowledge, by the thought that actions should not depend for their status as actions on anything external to them (including their causal history), and by a general scepticism about causation. In the preceding subsection, I commented briefly on Frankfurt's non-causalist position. Here I will consider Carl Ginet's (Ch. 5 and 1990).

In Chapter 1, Davidson issues the following challenge to non-causalists: given that when we act intentionally we act for reasons, provide an account of the reasons *for which* we act that does not treat (our having) those reasons as figuring in the causation of the relevant behaviour! The challenge is particularly acute when an agent has more than one reason for *A*-ing but *A*-s only for some subset of those reasons. For example, Al has a pair of reasons for mowing his lawn this morning. First, he wants to mow it this week and he believes that this morning is the most convenient time. Second, Al has an urge to repay his neighbour for the rude awakening he suffered recently when she turned on her mower at the crack of dawn and he believes that his mowing his lawn this morning would constitute suitable repayment. As it happens, Al mows his lawn this morning only for one of these reasons. In virtue of what is it true that he mowed it for this reason, and not for the other, if not that this reason (or his having it), and not the other, played a suitable causal role in his mowing?

In Chapter 5, Ginet develops an interesting response to the challenge in rebutting a pair of arguments against his position on free action.[24] Consider

---

Even so, his continuing so to intend was a causal sustainer of his actual continued coasting. That continued coasting was a willing coasting and an action. The counterfactual coasting sustained by paralysis is neither a willing coasting nor an action; it is not the same event as the continued coasting at the actual world. The continued coasting at the actual world is *motivated*—a causal notion—by Al's persisting intention to coast. If Al were to abandon the intention, *that* coasting would cease.

[24] For a related answer, see Wilson (1989), *Intentionality of Human Action*.

a 'reasons explanation' of the form '*S V*-ed in order (thereby) to *U*' (p. 118). 'The only thing *required* for the truth of a reasons explanation of this sort', Ginet writes, 'besides the occurrence of the explained action, is that the action have been *accompanied* by an intention with the right sort of content' (p. 119). In particular, it is not required that the intention figure in the causation of *V* or any part of *V*. 'Given that *S* did *V*', Ginet contends, it is sufficient 'for the truth of "*S V*-ed in order to *U*"' that 'concurrently with her action of *V*-ing, *S* intended by *that* action to *U* (*S* intended *of* that action that by it she would *U*).' He adds: 'If from its inception *S* intended of her action of opening the window that by performing it she would let in fresh air . . . , then *ipso facto* it was her purpose in that action to let in fresh air, she did it in order to let in fresh air.' Further, 'The content of the intention is . . . the proposition "By *this V*-ing (of which I am now aware) I shall *U*". It is owing to this direct reference that the intention is about, and thus explanatory of, *that particular* action' (p. 120).

Ginet's position is problematic.[25] Imagine that *S* had the following two *de re* intentions while opening the window, and that both were present at the time of the completion of that action: the intention, *N*, of her opening the window, 'that by it she would' let in some fresh air, and the intention, *O*, of her opening the window, that by it she would gain a better view of the street. Suppose that a neuroscientist, without altering the neural realization of *N* itself, renders that realization incapable of having any effect on *S*'s bodily movements (and any effect on what else *S* intends) while allowing the neural realization of *O* to figure normally in the production of movements involved in *S*'s opening the window. Here, one might plausibly argue, *O* helps to explain *S*'s opening the window and *N* does not. And if that is right, Ginet is wrong. For, on his view, the *mere presence* in the agent of an intention about her *V*-ing (where *V*-ing is an action) is sufficient for that intention's being explanatory of her action.

Ginet argues elsewhere that agents sometimes act in the absence of any relevant desire or intention whatever. He claims that some volitions (volitions being actions) are cases in point, as are some 'exertions' of the body. For example, 'a voluntary exertion could occur [owing to an associated volition] quite spontaneously, without being preceded or accompanied by any distinct state of desiring or intending even to try . . . to exert, and it would still be an action, a purely spontaneous one'.[26] In the case of a *voluntary* exertion of the body, Ginet says, 'a causal connection between the willing and the body's exertion is required',[27] but the volition itself, for Ginet, need not be caused (even in part) by, or concurrent with, any desire or intention.

So suppose that in *S*, standing within arm's reach of a window, a steady stream of volitions spontaneously springs up (volitions being momentary actions[28]), as a result of which *S*'s body moves in such a way as to come into contact with the window and smoothly open it in a conventional way. Suppose further that all this happens in the absence of any relevant intention or desire. Since the volitions produce the bodily movements that in turn cause the window to open, we have the makings of a causal explanation of all but the volitional element in *S*'s opening the window. (The first and spontaneous volition in the stream is the 'initial part or stage' of the voluntary exertion and the action.[29]) And the volitional element, on Ginet's view, needs no explanation at all.

Augment the case with an intention. Imagine that *S* intends of her opening the window 'that by it she' will let in some fresh air, but that her intention, *N*, is incapable of playing a causal role in producing the bodily movements or members of the volitional stream, owing again to the machinations of our neuroscientist. I do not see how *N* can have any more explanatory significance in the present case than it apparently had in the earlier *two*-intention case—that is, none at all. One might be tempted to think that the intention is explanatory of the action, on the grounds that, in the absence of *any* relevant intention or desire, *S*'s opening the window—that action—would be incomprehensible. But if Ginet is right, such an action requires no intention or desire at all for its occurrence: a spontaneous stream of volitions can do the work. Moreover, for readers who think it bizarre that, in the absence of any relevant intention, a steady stream of 'volitions' of a kind suitable for window-opening bodily movements would occur in an agent, and who therefore want to bring some intention into the explanatory picture, the best candidate would seem to be an intention that is *causally explanatory* of the supposed occurrence of the causally effective volitions.

I should add that, in my opinion, Ginet is entirely correct in rejecting the thesis that 'reasons explanation' requires the truth of determinism. But notice that causal explanation does not entail determinism, unless causation is essentially deterministic. Audi's position on reasons-explanation in Chapter 4 is in the causalist camp, but the essay displays no commitment to determinism.

---

[25] The following criticism of Ginet's view derives from Mele (1992*a*), *Springs of Action*, ch. 13.

[26] Ginet (1990), *On Action*, 9.

[27] Ibid. 39.

[28] Ibid. 32–3.

[29] Ibid. 30.

(F) *Agent causation*

A theory of action that would merit significant attention in a volume on free will, but is not represented in this volume, is the 'agent causation' theory. Although I cannot do it justice here, a brief description is in order. Some theorists have viewed free will or free action as incompatible with *caused* willing or acting, on any standard construal of causation. However, if freedom of these kinds requires agential control, and if control is, as it seems to be, a causal phenomenon, these theorists face a significant problem. Roderick Chisholm, Richard Taylor, and others have appealed to 'agent causation' in this connection. As Taylor describes it, it is 'causation of events by beings or substances that are not events'.[30] Chisholm claims, similarly, that we have agent causation only when 'there is some event, or set of events, that is caused, *not* by other events or states of affairs, but by the man himself, by the agent'.[31] 'On this view', Taylor writes, 'it is a man himself, and not merely some part of him or something within him'—for example, 'desires, choices, acts of will, and so on' (p. 49)—'that is the cause of his own activity' (p. 51). The view supposedly has the dual advantage for libertarians of circumventing the control problem, 'by conceding that human behaviour is caused', and blocking determinism, by placing agents at the beginning of causal chains issuing in actions: 'Some . . . causal chains, on this view, have beginnings, and they begin with agents themselves' (Taylor 1963: 52; cf. Chisholm 1966, Clarke 1993,[32] Taylor 1966,[33] and Thorp 1980[34]) For instructive criticism of agent causation by a libertarian, see Kane 1989.[35]

(G) *Volitions and trying*

Some philosophers have argued that volitions, or acts of willing, are essential to intentional action (e.g. Ginet 1990; McCann 1974;[36] O'Shaughnessy, Chapter 3). In itself, this idea is neutral on the question whether causalism is correct. Volitionists may seek to supplement non-volitionist causal theories of action with volitions, or they may defend an alternative to causal theories.

Brian O'Shaughnessy, in Chapter 3, identifies willing (or volition) with *trying* (p. 54). Trying, for O'Shaughnessy, is an 'inner' or 'mental' event. When one successfully tries to raise one's right arm, 'trying to raise the arm causes the act-neutral event of arm rising' (p. 62). 'The act of raising the arm is a complex event, constituted out of a causally linked pair of events, the trying and the arm rising' (p. 70). If an event's counting as a trying depends upon how it is caused, O'Shaughnessy's analysis of physical action is consistent with causalism about action.

Trying, O'Shaughnessy argues, serves a 'crucial bridge function between mind and body, not unlike that allotted by Descartes to the pineal gland' (p. 65). Trying 'is like a psychic promontory that all but juts into the physical world'. It is a 'mental event' that is '*standardly* a cause of physical change' (p. 66). Imagine (if you can) that an agent who is unknowingly paralysed tries to raise her right arm and there is no bodily motion at all. Then, we can say, her trying is a purely mental event, an act of will. Suppose that she subsequently recovers and again tries—this time successfully—to raise the arm. Should we say that her trying in this case, too, is a purely mental event, something that falls short of jutting into the physical world?

I doubt it (and O'Shaughnessy himself retracts the idea in *The Will*, ii. 100–11). Recently, I tried to raise my right arm in a contest of strength, while Kent, a large and powerful fellow, was trying to hold it down. I tried so long and so hard that I broke out in a sweat and strained a muscle. Would a purely mental trying have affected me so? Trying to *A* may be understood as making an effort to *A*, however modest that effort may be. (When, e.g. I intentionally turn my computer on, I am trying to do that, even if I encounter no special resistance and need to make no special effort.) The trying continues as long as the effort does. My effort to raise my arm included considerable muscular exertion. If my trying to raise my arm was my making that effort, my trying included this muscular exertion.

My trying was not *limited* to my muscular exertion, however. To see this, consider another attempted arm-raising. Although an injury has deprived Ben of all sensation in his right arm, he has retained the ability to move it. During an experiment, a physician blindfolds Ben, administers a dose of curare, and asks Ben to raise his right arm over his head and hold it there until he receives further instructions. Ben proximally intends to do so. The drug takes effect after signals from Ben's brain have started down the

---

[30] R. Taylor (1963), *Metaphysics* (Englewood Cliffs, NJ: Prentice-Hall), 52.

[31] R. Chisholm (1966), 'Freedom and Action', in Keith Lehrer (ed.), *Freedom and Determinism* (New York: Random House), 17.

[32] R. Clarke (1993) 'Towards a Credible Agent-Causal Account of Free Will', *Noûs* 27: 191–203.

[33] R. Taylor (1966), *Action and Purpose* (Englewood Cliffs, NJ: Prentice-Hall).

[34] J. Thorp (1980), *Free Will: A Defence Against Neurophysiological Determinism* (London: Routledge & Kegan Paul). J. Bishop (1983) in 'Agent Causation', *Mind* 92: 61–79, defends a related brand of agent-causation, motivated largely by problems that causal deviance poses for the project of providing an 'event-causal' analysis of intentional action. For such an analysis of intentional action that accommodates causal deviance, see Chapter 10.

[35] R. Kane (1989), 'Two Kinds of Incompatibilism', *Philosophy and Phenomenological Research* 50: 219–54.

[36] H. McCann (1974), 'Volition and Basic Action', *Philosophical Review* 83: 451–73.

试读结束：需要全本请在线购买： www.ertongbook.com

efferent pathway to his right arm, but paralysis sets in before the muscles can contract. Moments later, Ben believes that he has raised his arm and that he is now holding it over his head. When the blindfold is removed, he is surprised to see his right arm at his side. Ben sincerely reports that he tried to raise it and that he felt certain that he raised it.[37]

Ben's surprise is nicely explained on the hypothesis that he tried to raise his arm. While trying, Ben would receive a kind of feedback, 'efferent copy', registering that motor signals were sent. Receipt of this feedback, a product of Ben's trying, would explain his experience of trying. Since that experience is not accompanied by evidence of failure, Ben predictably would believe that he raised his arm.[38]

If Ben tried to raise his arm, his trying involved no muscular motion. In what did the trying consist? Frederick Adams and I have argued that it consists in a certain neurophysiological event with a certain causal history.[39] As we see it, Ben began trying to raise his arm when his proximal intention to raise it began to play its functional role. Certain neurophysiological effects of the acquisition of that intention—effects appropriate to his raising his arm—realize his trying to raise his arm partly in virtue of their causal history. Ben's attempt was both initiated by the acquisition of his proximal intention and motivationally sustained by the intention's continued presence. On this view, my own attempt to raise my arm in the contest with Kent encompassed neurophysiological events of the sort that occur in Ben's case and muscle contractions, as well.

3. INTENTION AND GENERAL PROBLEMS FOR ANALYSES OF INTENTIONAL ACTION

Intentional action is of primary importance in the philosophy of action. If there were no intentional actions, actions would be of little interest at best, and perhaps there would be no actions at all. (Davidson has argued that every action is intentional under some description,[40] a thesis Hornsby endorses in Chapter 12.) In discussions of freedom of action, intentional action occupies centre stage: we are much less concerned with conditions for the freedom of non-intentional actions. And although we are morally accountable for some unintentional actions, as in cases of negligence, moral assessment of actions is focused primarily upon intentional actions.

Presumably, there is some interesting connection between intentional actions and intentions. But what are intentions? And what is the connection? In this section, I discuss two competing approaches to characterizing intention and then turn to some general problems for the project of analysing intentional action.[41]

There is, predictably, considerable agreement that intentions are closely linked to desires and beliefs. It is generally recognized that intention has a motivational dimension, and 'desire' (like 'want') is often used in the literature as an umbrella term for motivation. Further, intention is widely, if not universally, regarded as involving a 'confidence' condition of some sort, a condition naturally explicated in terms of belief. Few people are inclined to maintain that a person who believes that her chances of winning today's lottery are about one in a million *intends* to win the lottery, no matter how strongly motivated she is to win. However, philosophers are divided on how close the connection is between intentions, on the one hand, and desires and beliefs, on the other. Some—attracted, perhaps, by the popular idea that desire and belief are the most fundamental representational states of mind[42]—hold that intentions are reducible to combinations of desires and beliefs.[43] Others have argued that attempts at such reduction are doomed to failure.[44]

In Chapter 6, Wayne Davis offers a reductive belief/desire analysis of intention:

> *S* intends that *p* iff [i.e. if and only if] *S* believes that *p* because he desires that *p* and believes his desire will motivate him to act in such a way that *p*. (p. 147)

Believing *p*, as Davis conceives it, 'is equivalent to being more certain of *p* than of not-*p*' (p. 133), and 'desiring *p* is equivalent to preferring *p* rather than not-*p*' (p. 135). (Davis's conception of desiring is at odds with a common conception according to which the same person can simultaneously desire to A and desire not to A: e.g. watching her son struggle with his homework, Ann may desire to help him with it; but she may also, at the same time, desire not to help him with it, thinking that he would benefit most if he were to complete his assignment on his own. Obviously, Ann cannot simultaneously prefer each of the options identified to the other.)

---

[37] This case, a version of which appears in F. Adams and A. Mele (1992), 'The Intention/Volition Debate', *Canadian Journal of Philosophy* 22: 323–38, is inspired by the description of Landry's patient quoted in W. James (1981), *The Principles of Psychology*, ii (Cambridge, Mass.: Harvard University Press), 1101–3.

[38] For further discussion, and references to some relevant empirical literature, see Adams and Mele (1992), 'The Intention/Volition Debate'.

[39] Ibid.

[40] Davidson (1980), *Essays on Actions and Events*, ch. 3.

[41] See Wilson (1989), *The Intentionality of Human Action* for a third approach, according to which intentions are not attitudes.

[42] For resistance to this idea, see Brand (1984), *Intending and Acting*, and J. Searle (1983), *Intentionality*.

[43] See, e.g. R. Audi (1973), 'Intending', *Journal of Philosophy* 70: 387–402, and id. 'Intending, Intentional Action, and Desire', in J. Marks (1986) (ed.), *The Ways of Desire* (Chicago: Precedent); M. C. Beardsley, 'Intending', in A. Goldman and J. Kim (1978) (eds.) *Values and Morals* (Dordrecht, Neth.: Reidel); W. Davis, Ch. 6.

[44] See, e.g. Brand (1984), *Intending and Acting*; M. Bratman (1987), *Intention, Plans, and Practical Reason* (Cambridge, Mass.: Harvard University Press), and id. Ch. 8; Davidson (1980), *Essays on Actions and Events*, ch. 5; G. Harman, Ch. 7; H. McCann (1986*a*), 'Rationality and the Range of Intention', *Midwest Studies in Philosophy* 10: 191–211; Mele (1992*a*), *Springs of Action*; and Searle (1983), *Intentionality*.

Although Davis's is the most promising belief/desire analysis of intention that I have seen, it is problematic.[45] Since, for Davis, desiring is preferring, two different senses of 'preference' merit attention. In one sense—an *evaluative* one—preferring A to B is giving A a higher evaluative rating than B. In another sense—a *motivational* one—preferring A to B is being more strongly motivated to A than to B.[46] Suppose that Ann satisfies the belief/desire conditions specified in Davis's analysis, on a reading of 'desire' as evaluative preference. Watching her son, Ann believes that she will refrain from helping him because she desires (= prefers in the evaluative sense) this and believes that her desire will motivate her to refrain from helping. Ann knows that, ordinarily, she acts on the basis of her evaluative preferences, and she deems it likely that she will do so in this case too: she is 'more certain' of this than of the contrary. Even so, given how pathetic her son looks, Ann is sorely tempted to help him, and she is unsettled about what to do. She is considering giving in to temptation. This seems coherent. If it is, Davis's analysis fails, on the current reading; for if Ann is unsettled about what to do and is considering helping her son, she does not—not yet, anyway—*intend* to refrain from helping him.

Essentially the same problem arises on the motivational reading. Suppose that although Ann believes that it would be best not to help her son, she is more strongly motivated to start helping him soon than to refrain from doing so. Suppose, further, that she knows she usually succumbs to temptation in cases of this kind and, accordingly, is 'more certain' that her desire to help will motivate her to help than that it will not so motivate her. However, Ann still is unsettled about whether to help. She is considering making an effort to resist temptation for the child's good. Given Ann's unsettledness, it is false that she intends to help her son—even though she satisfies Davis's conditions, on the current reading of 'desire'.[47]

---

[45] For an objection like the one to be advanced, see H. McCann (1986*b*), 'Intrinsic Intentionality', *Theory and Decision* 20: 251–2.

[46] For roughly this distinction, see G. Watson (1977), 'Skepticism about Weakness of Will', *Philosophical Review* 86: 320–1 and A. Mele (1987*a*), *Irrationality* (New York: Oxford University Press).

[47] Compare this case with Davis's case of Alan (p. 147). For detailed discussion of cases of this kind, see Mele (1992*a*), *Springs of Action*, ch. 9 and id. (1995), *Autonomous Agents*, ch. 3.

The points just made about settledness point toward an alternative view of intention as an attitude that is irreducible to belief/desire complexes. Functions plausibly attributed to intentions include initiating and sustaining intentional actions, guiding intentional behaviour, helping to coordinate agents' behaviour over time and their interaction with other agents, and prompting and appropriately terminating practical reasoning. Some philosophers have advanced non-reductive accounts of intention designed to accommodate these functions (see n. 44). According to a representative account of this kind,[48] intentions are executive attitudes toward plans. Plans—which range from simple representations of 'basic' actions to complex strategies for achieving remote goals—constitute the representational contents of intentions.[49] What distinguishes intentions from other practical attitudes—e.g. desires—is their distinctive practical nature. Although one can harbour a desire to do something without being at all settled upon doing it, to intend to do something is, in part, to be settled upon doing it (but not necessarily irrevocably). Such settledness upon a course of action constitutes a psychological commitment to executing the pertinent plan of action, a commitment of a kind arguably constituted exclusively by intentions.[50]

As I mentioned, it is natural to suppose that intention and intentional action are importantly related. It is noteworthy in this connection that Robert Audi's account of acting for a reason in Chapter 4 makes no mention of intention; for acting for a reason is often *identified* with acting intentionally.[51] However, Audi, like Davis, attempts to analyse intentions as complexes of beliefs and desires,[52] and he regards the reasons for which we act as, in part, states of affairs that express something we desire and are connected to action through belief (Ch. 4: 76). If intentions can be reduced to belief/desire complexes, an analysis of acting for a reason might *implicitly* accord intentions a major role in intentional action.

---

[48] Mele (1992*a*), *Springs of Action*.

[49] Roughly speaking, *basic* actions differ from non-basic actions in not being performed by way of the agent's performing another action.

[50] The commitment aspect of intention receives detailed treatment in Bratman (1987), *Intentions, Plans, and Practical Reason*. Also see R. Audi (1991), 'Intention, Cognitive Commitment, and Planning', *Synthese* 86: 361–78. Incidentally, I take decisions (when construed as states of mind) to be actively produced intentions. In deciding, one forms an intention; the intention formed may be termed a 'decision' (Mele (1992*a*), *Springs of Action*, ch. 9). But not all intentions are actively produced (Mele (1992*a*), 141).

[51] Taking off from Gregory Kavka's (1983) toxin puzzle ('The Toxin Puzzle', *Analysis* 43: 33–6), I have argued both that counterexamples to this alleged identification are provided by some bizarre scenarios and that the bizarreness required to generate a telling counterexample is itself revealing (A. Mele (1992*c*), 'Intentions, Reasons, and Beliefs: Morals of the Toxin Puzzle', *Philosophical Studies* 68: 171–94).

[52] Audi (1973), 'Intending', and id. (1986) 'Intending, Intentional Action, and Desire'.

For expository purposes, it will be useful to have before us a pair of proto-analyses of doing something intentionally, one framed in terms of intentions and the other in terms of reasons. With the stipulation that '*A*' is an action variable, the following will do:

(A1) *S* intentionally *A*-ed if and only if *S* *A*-ed in the way that *S* intended to *A*.
(A2) *S* intentionally *A*-ed if and only if *S* *A*-ed for a reason.

Both proto-analyses enjoy intuitive support. Again, there is, presumably, some important, substantial connection between what we intend to do and what we do intentionally. Similarly, it seems, intentional action is a species of behaviour intimately bound up with agents' desires and beliefs; and desires, perhaps typically in conjunction with beliefs linking desired goals to prospective instrumental behaviour, arguably constitute *reasons* for action. I will consider four problems for the proto-analyses. Both the problems and the proto-analyses are neutral on issues that divide causalists and noncausalists about intentional action.

(A) *Side effects*

Consider an example of Gilbert Harman's (Ch. 7: 151). 'In firing his gun', a sniper who is trying to kill a soldier, 'knowingly alerts the enemy to his presence'. Harman claims that although the sniper 'does not intend to alert the enemy to his presence', he nevertheless *intentionally* alerts the enemy, 'thinking that the gain is worth the possible cost'. If Harman is right, *A1* and *A2* are both false. Not only does the sniper not intend to alert the enemy, he does not alert them *for a reason* either (even if his alerting them is part of some 'larger' action that is done for a reason). Michael Bratman makes the same general claim, illustrated by a scenario featuring a runner who reluctantly wears down some heirloom shoes (Ch. 8: 199–203).

Harman's sniper and Bratman's runner do not unknowingly, inadvertently, or accidentally perform the actions at issue. For that reason, many will deny that the sniper *unintentionally* alerted the enemy and that the runner *unintentionally* wore down his shoes. But that denial does not, in any *obvious* way, commit one to insisting that the actions in question are *intentional*. Perhaps there is a middle ground between intentional and unintentional action. Arguably, actions that an agent in no way aims at performing but that are not performed unknowingly, inadvertently, or accidentally are properly located on that middle ground. They might be *non*-intentional, as opposed to *un*intentional (cf. Ch. 10: 230–31).

(B) *Belief-constraints*

Anticipated side effects are not the only alleged problem for the thesis—dubbed by Bratman the 'Simple View'—that intentionally *A*-ing entails intending to *A*. Some putative belief-constraints on intentions, or *rational* intentions, also pose problems. Bratman argues that intention has a normative side that includes, among other things, demands that an agent's intentions be internally consistent (individually and collectively), consistent with the agent's beliefs, and means-end coherent (Ch. 8; cf. Harman, Ch. 7). *Rational* intentions, he maintains, meet those demands. Concerning beliefs in particular, he contends that *S* *rationally* intends to *A* only if, 'other things being equal', *S* does 'not have beliefs inconsistent with the belief that [he] will *A*' (p. 186).

The normative demands figure significantly in a much-discussed argument of Bratman's against the Simple View (Ch. 8). The argument turns on an example featuring a pair of video games and an ambidextrous player whom I shall call Bart. Bart's task is to hit video targets with video missiles. In the main case, he is simultaneously playing two games, each with its own target and firing mechanism, and he knows that the machines are 'so linked that it is impossible to hit both targets' (p. 184). (He knows that hitting a target ends both games, and that 'if both targets are about to be hit simultaneously', both machines shut down before the targets can be hit.) Bart simultaneously tries to hit the target on machine 1 and tries to hit the target on machine 2. He succeeds in hitting the former—'in just the way that [he] was trying to hit it, and in a way which depends heavily on [his] considerable skills' (p. 184)—but, of course, he misses the latter.

Supposing that Bart hits target 1 *intentionally*, proponents of the Simple View must say that he intended to hit it. Since Bart's attitude toward hitting that target is not relevantly different from his attitude toward hitting target 2, they apparently must hold as well that he intended to hit target 2. Bratman claims that having *both* intentions, given what Bart knows (namely, that he cannot hit both targets), would be irrational. Yet, *it seems perfectly rational of Bart to have proceeded as he did.* So given the point about the symmetry of Bart's attitudes toward the targets, Bratman concludes that he did not have either intention. And if Bart hit target 1 intentionally in the absence of an intention to hit it, the Simple View is false.

If trying to *A* requires intending to *A*, Bratman's video games argument fails. For then, trying to hit each target, Bart would intend to hit each. Hugh McCann has argued that trying to *A* and intending to try to *A* are each sufficient for intending to *A*.[53] And I have argued that his arguments are unsuccessful.[54] Here, I set the details of that dispute aside and take up a portion of McCann's defence of the Simple View in Chapter 9.

Some critics of the Simple View are also critical of the idea that intentions are reducible to complexes of beliefs and desires (e.g. Bratman, Harman, and Mele), and McCann argues that they are in danger of having to settle for an unwanted reductive analysis of intention (Ch. 9). Bratman, who suggests that a 'guiding desire'—e.g. to hit target 1—can play the role of an intention,[55] receives the brunt of the attack. McCann observes that once it is conceded that desires can stand in for intentions, reductionists will justifiably ask what functional need there is for a notion of intention that is irreducible to desire and belief. However, opponents of the Simple View need not follow Bratman in appealing to guiding desires. On my own view, for example, intentions to *try* to *A* can stand in for intentions to *A*, but intentions to try to *A* are *intentions*.[56] The agent's attitude toward *A*-ing is not one of intending *to A*, but neither is it merely one of desiring to *A*. It is, rather, an intending-to-try attitude toward *A*-ing, and intending to try is a species of intending. (My view is a version of Bratman's 'Single Phenomenon View' (Ch. 8: 194–203).)

Normally, at least, one who intends to try to *A* has *A*-ing as a goal, purpose, or objective. McCann contends that 'there is no ordinary sense in which terms like 'goal' or 'purpose' signify objectives that guide deliberation and behaviour, but fall short of being intentions' and that an intention to *A* is 'implicit in' an intention to try to *A* (Ch. 9: 221). I disagree. Poor Lydia, who has only one dollar, would love to have a million. There are no lotteries in her state, but there is a weekly million-dollar contest for amateur golfers. Contestants pay a dollar for the privilege of taking one shot at making a hole in one from a distance of 180 yards. Lydia has never hit a golf ball, but desperately wanting to become a millionaire and thinking that there is a remote chance that she will make a hole in one, she enters the contest. She has seen golf on television, and she estimates her chances of holing her shot at about one in a million. As Lydia eyes the ball, she deliberates about how she might achieve the *goal* or *objective* of making a hole in one, giving special attention to what club to use. She selects a three wood, lines up the shot, and then swings hard, with the goal or objective of making a hole in one. Lydia does not hit the ball just for the sake of hitting it. Nor is her objective in hitting it limited to something less than hitting a hole in one. Her goal is to hit a hole in one, thereby winning a million dollars.

This little story evidently is coherent and completely in line with ordinary usage of such terms as 'goal' and 'objective'. So McCann must hold that, despite her awareness of the astronomical odds against her holing the shot, Lydia *intends* to hole it, whereas others would contend that, instead, she fervently hopes or strongly desires to hole it, or intends to try to hole it. Since, as McCann says, ordinary senses of terms are at issue and not philosophers' technical uses, the assertion that Lydia intends to hit a hole in one seems seriously mistaken (cf. Davis, Ch. 6: 133, and Harman, Ch. 7: 151 on winning lotteries). (Put yourself in Lydia's shoes as she approaches the ball. Given your belief that your chances of holing the ball are about one in a million, would it be true that you intend to hole it, or would it rather be the case that you intend to try to hole it?) Further, it is utterly plausible that Lydia intends to *try* to hit a hole in one. Thus, it is plausible that she has an *intention* regarding her goal of hitting a hole in one, even though that intention is not an intention *to hit a hole in one*.

The objection just advanced is directed at an important element in McCann's attempted refutation of some criticisms of the Simple View. It is not an objection to the Simple View itself. Ultimately, I think, the fate of the Simple View rests on whether its truth is required for the *explanation* of intentional actions. I have argued elsewhere that the truth of the Simple View is not required for this purpose.[57] But I have argued as well that the belief-constraints to which McCann objects are not required for this purpose either.[58]

(C) *Intrinsically motivated actions*

I turn now to reasons and to *A2*. On a popular account, the reasons for which we act—*effective reasons*—are complexes of beliefs and desires or pro-attitudes (Davidson, Ch. 1). Thus, the reason for which I crossed the road might be constituted by a desire to get to the other side and a belief that doing so requires a crossing. This account of effective reasons seems not to do justice to what Audi (Ch. 4) terms 'intrinsically motivated actions'—actions done for their own sakes, from 'intrinsic' desires. When something, *A*, is done for its own sake *alone*, it is *not* done from a desire for something further, *F*, and a belief that identifies *A*-ing as suitably related to *F*. However, if the reasons for which we *A* are sometimes constituted wholly by intrinsic desires to *A*, such cases can easily be handled.[59] Intrinsically motivated actions arguably are a problem, not for *A2* itself, but for a particular conception of effective reasons.

---

[53] H. McCann (1986*a*), 'Rationality and the Range of Intention'; and id. (1989), 'Intending and Planning: A Reply to Mele', *Philosophical Studies* 55: 107–10; cf. Ch. 9 and F. Adams (1986), 'Intention and Intentional Action: The Simple View', *Mind and Language* 1: 281–301.

[54] A. Mele (1989), 'She Intends to Try', *Philosophical Studies* 54: 101–6; and id. (1992*a*), *Springs of Action*, 132–5.

[55] Bratman (1987), *Intention, Plans, and Practical Reason*, 137; cf. Ch. 8.

[56] Mele (1992*a*), *Springs of Action*, ch. 8.

[57] Ibid.

[58] My strategy (Mele (1992*a*), *Springs of Action*, ch. 8) was to catalogue various functions of intentions and to argue that intention's having these functions does not depend upon the truth of various alleged belief-constraints.

Rosalind Hursthouse (1991)[60] appeals to a species of intrinsically motivated action, 'arational action', in an attempt to undermine *A2*. Examples of arational actions include striking an inanimate object in anger and gouging out the eyes in a photograph of a hated person. She also adduces, but in another category, 'actions prompted by odd physical cravings'—e.g. licking something furry when 'seized by a sudden desire' to do so (pp. 62–3). Such actions, as Hursthouse observes, often are not done for the sake of some further goal, and they typically seem unreasonable. But it would be a mistake to infer from this that they are done for no reason at all. If our reasons can be every bit as bizarre as our actions, proponents of *A2* have no special cause for worry. A man with an irrational urge to drink a can of paint (Davidson, Ch. 1) and the knowledge that drinking the paint requires removing the lid might *pry off the lid* for a reason, and I have not encountered compelling grounds for thinking that he cannot *drink the paint* for a reason, too—a reason constituted by an intrinsic desire to drink it.

(D) *Luck*

A more interesting problem for *A2* has attracted little attention (as a problem for *that* thesis). In Chapter 4, Audi claims that all actions done for a reason are intentional (p. 104). This popular thesis is challenged by some cases of extraordinary luck. Connie, who has never fired a gun, is offered a large cash prize for hitting the bull's-eye on a distant target that even experts normally miss. She carefully aims and fires, hitting the target dead centre in just the (direct) way she hoped she would. Many readers, I think, would happily (but perhaps mistakenly) say that Connie's hitting the bull's-eye—that action—was done for a reason.[61] After all, she wanted the money and believed that to get it she must hit the bull's-eye, and this helps to explain her carefully aiming and firing at the target. (That Connie hit the bull's-eye for a reason is compatible with Audi's account of acting for reasons, but see Chapter 4: 89 for an indication that he might reject this assessment of Connie's action.) But was Connie's hitting the bull's-eye an *intentional* action? To simplify matters, suppose that Connie has no natural talent for marksmanship: she tries equally hard to win even larger prizes for duplicating the feat, fires five hundred rounds at the target, and does not even come close.

Here intuitions differ. According to Christopher Peacocke, an agent who makes a successful attempt 'to hit a croquet ball through a distant hoop' *intentionally* hits the ball through the hoop.[62] But Brian O'Shaughnessy maintains that a novice who similarly succeeds in hitting the bull's-eye on a dartboard does not intentionally hit the bull's-eye.[63]

Luck is also a problem for *A1*, of course. Just suppose that Connie, who mistakenly thinks that modern weaponry makes target shooting easy, *intends* to hit the bull's-eye by aiming and firing at it. She hits it in just the way intended, but was her hitting it an intentional action? Readers inclined to answer affirmatively should consider a similarly benighted person who intends to disarm a doomsday device. She thinks that all she need do to disarm it is to punch in any ten-digit code, whereas, in fact, only one ten-digit code will work; and wanting to disarm the machine, she intends to disarm it by entering ten digits. If she luckily punches in the right code, thereby disarming the machine, does she disarm it intentionally? (Does she disarm it *for a reason*—perhaps one constituted by a desire to save the world and a belief that she can ensure her doing that by disarming the machine?) The problems that luck poses for a proper understanding of intentional action are examined in Chapter 10.[64]

4. CLOSING COMMENTS

My aim in the essay was to introduce readers both to central issues and debates in the philosophy of action and to the twelve essays that follow. The objections I have voiced are intended primarily to promote reflection on the following essays, not to settle issues. A comprehensive philosophy of action will include a stand on each of the main issues discussed here: the nature of action, trying, action-explanation, reasons for action, intention, and intentional action. However, it should not be thought that a particular stand on one topic will commit one to a position on all the others. For example, causalists and non-causalists alike can accept or reject the idea that intentions (or effective reasons for action) are identical with belief/desire complexes, the thesis that trying is always a strictly mental action, and the Simple View of the connection between intention and intentional action. Students are encouraged, therefore, to investigate each issue raised in the following essays on its own terms.

---

[59] Ibid. ch. 6.

[60] R. Hursthouse (1991), 'Arational Actions', *Journal of Philosophy* 88: 57–68.

[61] For resistance, see A. Mele (1992*b*), 'Acting for Reasons and Acting Intentionally', *Pacific Philosophical Quarterly* 73: 355–74. Incidentally, intuitions tend to shift when the agent succeeds in a way that diverges significantly from her plan, or from her (perhaps tacit) assumptions about what a successful *A*-ing would involve. If the bullet had ricocheted off several rocks into the bull's-eye, would Connie's hitting the bull's-eye have been done for a reason?

[62] C. Peacocke (1985), 'Intention and *Akrasia*', in B. Vermazen and M. Hintikka (eds.), *Essays on Davidson* (Oxford: Clarendon Press).

[63] B. O'Shaughnessy (1980), *The Will*, ii (Cambridge: Cambridge University Press), 325; cf. Harman, Ch. 7, 151.

[64] Other apparent problems are posed by sudden and impulsive actions and by subsidiary actions. For a response, see Chapter 10.

My aim in selecting articles for this volume was to provide students with first-rate, accessible essays on traditional issues in the philosophy of action that have become dominant issues in the area. Thus, traditional issues in which interest has waned (e.g. action-individuation and basic action) receive only incidental attention in this volume, and less traditional topics in which interest is growing (e.g. group action and group intentions) are not represented here. Owing to reasonable constraints on space, much excellent work could not be included, but my Further Reading section provides some guidance.

Parts of this introduction derive from my (1992*d*) 'Recent Work on Intentional Action', *American Philosophical Quarterly* 29: 199–217. I am grateful to the executive editor for permission to use material from that article, to Michael Bratman and Peter Momtchiloff for helpful written comments on a draft of this introduction, and to Robert Audi, Michael Bratman, Peter Momtchiloff, and Paul Moser for advice about the selection of essays.

1

ACTIONS, REASONS, AND CAUSES

DONALD DAVIDSON

What is the relation between a reason and an action when the reason explains the action by giving the agent's reason for doing what he did? We may call such explanations *rationalizations*, and say that the reason *rationalizes* the action.

In this paper I want to defend the ancient—and common-sense—position that rationalization is a species of ordinary causal explanation. The defence no doubt requires some redeployment, but not more or less complete abandonment of the position, as urged by many recent writers.[1]

I

A reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action—some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable. We cannot explain why someone did what he did simply by saying the particular action appealed to him; we must indicate what it was about the action that appealed. Whenever someone does something for a reason, therefore, he can be characterized as (*a*) having some sort of pro attitude toward actions of a certain kind, and (*b*) believing (or knowing, perceiving, noticing, remembering) that his action is of that kind. Under (*a*) are to be included desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles,

Donald Davidson, 'Actions, Reasons, and Causes', *Journal of Philosophy* 60 (1963), 685–700.
© 1963 Donald Davidson. Reprinted by permission of the author.

---

[1] Some examples: G. E. M. Anscombe (1959), *Intention* (Oxford); Stuart Hampshire (1959), *Thought and Action* (London); H. L. A. Hart and A. M. Honoré (1959), *Causation in the Law* (Oxford); William Dray (1957), *Laws and Explanation in History* (Oxford); and most of the books in the series edited by R. F. Holland, *Studies in Philosophical Psychology*, including Anthony Kenny (1963), *Action, Emotion and Will* (London), and A. I. Melden (1961), *Free Action* (London). Page references in parentheses will all be to these works.

试读结束：需要全本请在线购买： www.ertongbook.com