ROUTLEDGE

Sarah K. Paul

PHILOSOPHY OF ACTION

A Contemporary Introduction

ROUTLEDGE CONTEMPORARY INTRODUCTIONS TO PHILOSOPHY

Philosophy of Action

This book offers an accessible and inclusive overview of the major debates in the philosophy of action. It covers the distinct approaches taken by Donald Davidson, G.E.M. Anscombe, and numerous others to answering questions like "what are intentional actions?" and "how do reasons explain actions?" Further topics include intention, practical knowledge, weakness and strength of will, self-governance, and collective agency. With introductions, conclusions, and annotated suggested reading lists for each of the ten chapters, it is an ideal introduction for advanced undergraduates as well as any philosopher seeking a primer on these issues.

Sarah K. Paul is Associate Professor of Philosophy at New York University Abu Dhabi and Global Network Associate Professor of Philosophy at New York University.

Routledge Contemporary Introductions to Philosophy
Series editor: Paul K. Moser, Loyola University of Chicago

This innovative, well-structured series is for students who have already done an introductory course in philosophy. Each book introduces a core general subject in contemporary philosophy and offers students an accessible but substantial transition from introductory to higher-level college work in that subject. The series is accessible to non-specialists and each book clearly motivates and expounds the problems and positions introduced. An orientating chapter briefly introduces its topic and reminds readers of any crucial material they need to have retained from a typical introductory course. Considerable attention is given to explaining the central philosophical problems of a subject and the main competing solutions and arguments for those solutions. The primary aim is to educate students in the main problems, positions and arguments of contemporary philosophy rather than to convince students of a single position.

Recently Published Volumes:

Philosophy of Science
4th Edition
Alex Rosenberg and Lee McIntyre

Philosophy of Western Music
Andrew Kania

Phenomenology
Walter Hopp

Philosophical Logic
John MacFarlane

Philosophy of Action
Sarah K. Paul

For a full list of published *Routledge Contemporary Introductions to Philosophy*, please visit: www.routledge.com/Routledge-Contemporary-Introductions-to-Philosophy/book-series/SE0111

Philosophy of Action

A Contemporary Introduction

Sarah K. Paul

ROUTLEDGE
Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2021
by Routledge
52 Vanderbilt Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2021 Taylor & Francis

The right of Sarah K. Paul to be identified as author of this work has been asserted by her in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data
A catalog record for this title has been requested

ISBN: 978-1-138-64273-7 (hbk)
ISBN: 978-1-138-64274-4 (pbk)
ISBN: 978-1-315-62977-3 (ebk)

Typeset in Times New Roman
by River Editorial Ltd, Devon, UK

Contents

Acknowledgments ix

1 Introduction: What Is the Philosophy of Action? 1

2 What Is the Problem of Action? 8
2.1 Activity and Passivity 9
2.2 Goal-Directedness 9
2.3 Attributability 10
2.4 "Actish" Phenomenal Quality 11
2.5 Voluntary Action 11
2.6 Rational Action, or Acting for Reasons 12
2.7 Practical Knowledge 12
2.8 Intentional Action 13
2.9 Intention 13
2.10 Autonomy, Identification, and Self-Governance 14
2.11 Further Choice Points 14
2.11a Which Cases Are Paradigmatic? 14
2.11b Questions About Action: Conceptual or Ontological? 15
2.12 Conclusion 15
Suggested Reading 16

3 Action Explanation 17
3.1 Guises of Rationalizing Explanation 18
3.2 Reasons for Action: Motivating vs. Normative 19
3.3 More on the "Why?" Question 20
3.4 Action Explanation: Four Views 22
3.4a The Rational Interpretation View 22
3.4b The Causal Theory of Action Explanation 24
3.4c Teleological Realism 29
3.4d Naïve Action Theory 33
3.5 Arational Action 35
Summary 36
Suggested Reading 37

4 The Ontology of Action 38
4.1 Which Things in the World Can Be Actions? 39
4.2 Under a Description 41
4.3 Basic Actions 41
4.3a Bodily Movements 43
4.3b Volitions 44
4.3c Beyond the Body 45
4.4 The Accordion Effect 46
4.5 How Many Actions? 47
4.6 The Causal Theory of Action 48
4.6a Objection 1: Deviant Causal Chains, Redux 53
4.6b Objection 2: The Disappearing Agent 56
4.7 Alternatives to the Causal Theory 57
4.7a Quietism 57
4.7b Agent-Causation and Causal Powers 58
4.7c Formal Causation 60
4.7d An "Actish" Phenomenal Quality 61
4.8 Omissions 62
4.9 Mental Actions 63
Summary 64
Suggested Reading 65

5 Intention 67
5.1 Methodological Priority: Present or Future? 67
5.2 Goal States and Plan States 69
5.3 Reductive Accounts of Intention 69
5.3a Predominant Desire 71
5.3b Predominant Desire Plus Belief 71
5.3c Evaluative Judgment 73
5.4 Plan States and Plan Rationality 74
5.5 Cognitivism About Intention 78
5.6 A Distinctively Practical Attitude 80
5.7 Intending and Intentional Action 82
Summary 84
Suggested Reading 85

6 Practical Knowledge 86
6.1 What Do We Mean by "Practical Knowledge?" 86
6.1a Knowledge Without Observation 87
6.1b Knowledge Without Inference 88
6.1c Mistakes Are in the Performance, Not the Judgment 88
6.1d The Cause of What It Understands 89
6.1e Contradicted by Interference 89
6.2 The Scope and Object of Practical Knowledge 90

6.3 Accounts of Practical Knowledge 93
6.3a Cognitivism About Intention 93
6.3b Imperfective Knowledge 95
6.3c The Inferential Account 97
Summary 99
Suggested Reading 100

7 Does Action Have a Constitutive Aim? 101
7.1 The Guise of the Good 102
7.2 The Aim of Self-Understanding 105
7.3 The Aim of Self-Constitution 107
7.4 The Will to Power 109
7.5 No Constitutive Aim 110
7.6 Implications for Ethics and Metaethics 110
Summary 112
Suggested Reading 113

8 Identification and Self-Governance 114
8.1 Frankfurt on Identification 115
8.2 Watson's Objection and Platonic Alternative 116
8.3 Frankfurt Redux: Wholeheartedness 118
8.4 Bratman on Self-governing Policies 119
8.5 Skepticism About Self-Governance: A Genealogical Worry 120
8.6 Self-Governance and Plan Rationality 121
Summary 123
Suggested Reading 123

9 Temptation, Weakness, and Strength of Will 125
9.1 Is Synchronic Akrasia Even Possible? 127
9.2 A Failure of Reasoning? 128
9.3 A Divergence Between Evaluation and Motivation? 130
9.4 Is Akrasia Necessarily Irrational? 131
9.5 Weakness of Will Over Time 133
9.6 Self-Control 135
Summary 139
Suggested Reading 140

10 Collective Agency 141
10.1 Questions and Constraints 141
10.2 Group Agents 142
10.3 Collective Intentions 144
10.3a Tuomela and Miller 145
10.3b Searle 146
10.3c Bratman 146

10.3d Velleman 148
10.3e Gilbert 149
10.4 Acting Together 150
Summary 152
Suggested Reading 153

11 Concluding Thoughts 154

Bibliography *157*
Index *164*

Acknowledgments

When I first arrived at graduate school to study philosophy, I didn't know what the philosophy of action was. I certainly didn't go with the intention of making any kind of intensive study of it. That I ended up becoming captivated by the topic and writing a dissertation on it is entirely due to the boundless enthusiasm and inexhaustible patience of Michael Bratman. These debates came alive when I saw them through his eyes, and as did the idea that I might one day be able to contribute something to the discussion. Much of what is in this book I learned from Michael, though he cannot be blamed for my errors and oversights. I am profoundly grateful for his support over the years, and fundamentally this book is for him.

I was subsequently welcomed into the community of philosophers of action by many people, but there are some I'd like specially to thank. Kieran Setiya, Luca Ferrero, Sergio Tenenbaum, and Sarah Buss all in various ways made me feel included and part of the conversation. Matthias Haase invited me to be a part of a "Netzwerk" of people interested in agency from all over Europe as well as the U.S. Participating in those meetings introduced me to a lot of wonderful people and greatly broadened my understanding of the variety of approaches one could take to these questions. And Matty Silverstein did the same by including me in his terrific series of workshops on "Normativity and Practical Reasoning."

My first tenure-track job was at the University of Wisconsin – Madison, and I owe that institution and my dear colleagues there a great deal. I agreed to write this book largely because Larry Shapiro told me to, and as my mentor, his thoughtful advice never erred. As my friend, I don't know about the advice, but Larry and Steve Nadler tolerantly let me air my frustrations with the writing process during our long runs around the arboretum. Elliott Sober, Russ Shafer-Landau, Mike Titelbaum, Alan Sidelle, and James Messina were also sources of great help and support. Much of this book was written with the aid of very generous research funding from the University of Wisconsin, including the Vilas Associates Award and Vilas Early Career Investigator Award.

Many thanks are also due to the students who took my undergraduate courses and graduate seminars in the philosophy of action over the years.

There are too many to mention, but their creative examples and acute observations infuse what I have written here. Special thanks to Megan Fritts, Camila Hernandez Flowerman, and Ben Schwan for help with some of the research that contributed to this book.

The chapter on self-control was written with the support of a grant from the Philosophy and Science of Self-Control project directed by Al Mele and funded by the Templeton Foundation. I learned a great deal about the topic from the other researchers who were part of that project, and I'm grateful to Al for the support. The grant was for work done in collaboration with Jennifer Morton, who has been my dearest friend, sounding-board, advocate, and critic since we met in graduate school many years ago. The book was finally finished while at New York University Abu Dhabi, and I'm very grateful for the generous research funding and great colleagues that I've found here.

Thanks also to my parents, Lowell and Sheila, and my husband, Bas. Bas is a true partner and co-parent without whose help and perspective I couldn't do much of anything. I can't say my daughter Ivy specifically helped much with finishing this book, but the delay was completely worth it. I'm grateful to Aneena Noor, without whose help with childcare I could not have finished this book during a pandemic. Thanks also to my editor, Andy Beck, for his patience with me when the book took much longer to write than I thought it would.

Finally, for specific comments on the manuscript, I'm deeply indebted to David Velleman, Kieran Setiya, Sergio Tenenbaum, Luca Ferrero, Matty Silverstein, Mikayla Kelly, Michael Bratman, Jennifer Morton, Larry Shapiro, Sarah Buss, and (especially) an anonymous reader for Routledge. This book is far better because of their generosity.

1 Introduction

What Is the Philosophy of Action?

In a bizarre incident that became the subject matter of *Palsgraf v. Long Island Railroad Co.*, a man is running to catch a departing train. A platform guard shoves him toward the train from behind while a member of the train's crew pulls him into the car. In the process, the man drops the package he is carrying. The package turns out to contain fireworks, which explode and cause a heavy scale 10 feet away to topple over onto Helen Palsgraf, a woman who is standing on the platform. She suffers injuries as a result.

This case involves a number of people doing things, both intentionally and by accident. To understand what happened, as the court must try to do in adjudicating the suit, a variety of questions must be answered. Why did the platform guard shove the traveler? Was his intention to help him catch the train, or did he do it out of malice? Did the traveler drop the package on purpose, expecting it to explode, or did he merely lose his grip on it? Is it right to say that injuring Palsgraf was something the platform guard did, or something the traveler did? Or was her injury nobody's doing, and simply the result of bad luck? The philosophy of action aims to understand precisely what questions like these are asking and how we might go about answering them.

To isolate exactly what we are interested in, it will be helpful to first explain what the topic is *not*, at least for the purposes of this book. First, the problem of action is not the same as the ethical problem of how we ought to act, or what we should be blamed and praised for. *Palsgraf* certainly raises questions about whether the guard and the traveler acted as they should have and whether they are to blame for the injuries they caused. These, however, are separate subject matters. Our first task is to identify what it is we are referring to when we ask whether someone *acted* ethically, or what she ought to *do*. As G.E.M. Anscombe famously wrote, "… it is not profitable for us at present to do moral philosophy; that should be laid aside at any rate until we have an adequate philosophy of psychology …" (1958, 1). It may turn out in the end that ethics and the philosophy of action are intertwined, perhaps because rational agency consists in pursuing the Good or in doing what one takes to be morally required. But such insights should be the conclusion of our investigation and not premises in it.

Second, Anscombe's admonition that we do philosophy of psychology might wrongly suggest that we are interested in the so-called "mind–body problem" or the problem of mental causation. These labels refer to thorny issues about how to understand the relationship between the mind, revealed to us from within as consisting of conscious thoughts and experiences, and the physical body, as revealed to us by empirical investigation. How could the conscious experience of seeing orange be nothing more than some neural activity in the brain? And how could it not be, if having such experiences can have bodily consequences such as uttering the sentence "What a beautiful shade of orange!" More generally, it seems that we can give physical explanations and mental explanations of the very same bodily event. When the guard's arm stretches out toward the traveler, we can explain what happened in terms of neural signals and muscle contractions, or we can clarify that he wanted to help the traveler and believed that pushing him at that moment was the way to do it. How, if at all, can these two kinds of explanations be reconciled?

The answers to these questions will be a crucial part of the complete story of what happens when an embodied agent moves herself. But before we can assess whether there is actually a direct conflict between physical and mental explanations, we must answer the question of whether an agent's raising his arm is really the same thing as an arm's rising. That is, are actions identical to bodily movements, or are they in a different category altogether? If they are not identical, then perhaps the two explanatory frameworks are not in conflict. Further, even if we solved the problem of understanding how the mind is related to physical states of the world, we would still face the puzzle of locating agency within that psycho-physical activity. Knowing exactly what a desire is in physical terms, and how desires cause the body to move, does not answer the question of whether a desire's causing a bodily movement amounts to an agent acting. For that, we need to do some philosophy of action.

Third, the problem of action is not the same as the problem of free will. As traditionally conceived, that question concerns whether we can be said to act freely if determinism is true, in the sense of freedom that many philosophers take to be required for moral responsibility. Can we ever be held accountable for what we do if our actions are preordained, either by the physical laws of nature or by the omniscience of a divine being? If the guard freely chose to push the traveler, and is legitimately blameworthy or praiseworthy for his choice or deed, must it have been genuinely possible for him to have chosen instead not to push? And if so, how could the existence of this alternative path be consistent with the rest of what we understand about the universe?

These questions are certainly relevant to our conception of ourselves and other beings as agents, and there is nothing wrong with using the label "philosophy of action" in a way that includes philosophical work on free will. However, this book will largely set such questions aside. On the narrower conception of the topic that is our focus, the central puzzles are both prior to and independent of our investigation into free will. They are prior because the question of whether our actions are free, and whether alternative possible

courses of action are ever available to us, presupposes a grasp of what an action *is*. They are independent because the interest is in what it is to act, whether or not it turns out that we have the capacity to act freely in the sense relevant to moral responsibility. In other words, even if we were to conclude that we have no free will, we would still be agents who spend our waking hours engaged in (unfree) actions of various kinds, and this would still be a phenomenon worth trying to understand.

Having trimmed back the hedges of surrounding issues, we are in a better position to think about the basic phenomenon of *doing* something. We will attempt to state the central questions more precisely in Chapter 2, but for now we can start with some intuitions. Think of all the things you have done just today. Perhaps you turned off your alarm, made coffee, spilled some on your shirt, arrived late to class, and spent some time fantasizing about lunch. "Doings" can be contrasted, first, with mere "happenings." Turning off your alarm was something that you did – you made it happen – but being awakened by it is something that happened to you. What exactly is the difference between making something happen and having something happen to you? Second, not everything you do is something you do intentionally, or as we more commonly say, "on purpose." Spilling coffee on your shirt was presumably not something that you meant to do, whereas making the coffee was intentional. Similarly, while arriving to class and arriving late to class seem to be one and the same event, we can assume that you intentionally went to class but did not intentionally arrive late. But what does it mean to do something intentionally or not, and how can the same event of your arriving to class at 9:23am be both an intentional action of yours and something you did unintentionally?

We assume that you intentionally went to class but did not intentionally arrive late because agents like us normally act in light of what we have good reason to do, or what it would make sense for us to do. There are many things to be said in favor of going to class, but not many in favor of being late. This suggests that whether or not your action was intentional is connected to whether you saw it as desirable, worth doing, or at least intelligible. That said, we do not always choose to do what we take to be the best thing to do. It seems perfectly possible to decide to sleep in and be late to class even though you sincerely believe that being punctual is more important than 15 more minutes of sleep. What on earth is going on when we intentionally do one thing while at the same time believing that another available option is substantially better? Another interesting feature of agents like us, and one that can be of use in combating this kind of weak-willed action, is that we do not always simply make up our minds in the moment about what to do. We sometimes deliberate in advance and plan ahead for what to do later, thereby forming intentions for the future. "Tonight," you might think, "I'll set my alarm 15 minutes earlier so that I'll be on time to class tomorrow."

Some of our actions are mental – fantasizing about lunch – while others involve far-reaching consequences in the non-mental world. For that matter,

some of the most consequential events we are involved in concern things that we intentionally *don't* do. Stanislav Petrov may have singlehandedly saved the world from all-out nuclear war when he refused to obey protocol and did not respond to what turned out to be a false alarm in the Soviet early warning system. And while singlehanded actions are often of particular interest to us, a great many of our actions are undertaken together with other people. We play dodgeball together, write books together, and participate in the institutions of society together. All these are the kinds of phenomena we will attempt to understand better in this book.

One way to motivate these puzzles about agency is to compare them to other well-known philosophical problems. For example, philosophy has long been occupied with investigating the nature of perception. We seek to better understand how the conscious mind is connected to the external world, and perception is one major way in which the mind interfaces with the world. We wrestle with questions like "How can we be in touch with objects that are external to our minds through the use of our senses?" and "Given the possibility of illusion, how can perception give us knowledge of reality?" Inspired by these questions, we might view action theory as a counterpart to the philosophy of perception. It is through acting that the mind interfaces with the world in the other direction, imposing itself on external reality rather than conforming itself to it. This framing will generate questions about action that tend toward the metaphysical, focusing on how agency fits together with the rest of the natural world.

A different contrast can be drawn between action and belief, thereby framing the topic in a way that will focus our attention on certain normative questions. Philosophers have long been interested in the question of why we should or shouldn't believe certain things – what counts as a sufficient reason to believe? We attempt to understand what it is to reason well about what to believe (call this "theoretical" reasoning), so that we can avoid certain tempting forms of logical error. Think of Sherlock Holmes reasoning about who might have committed the murder and forming the belief that it was a one-legged man and his small accomplice, for the reason that the tracks in the dust indicate these unusual physical features. Viewed through this lens, the state or activity of believing is defined partly in terms of its being the conclusion of theoretical reasoning, or a way of responding to these kinds of reasons. Likewise, we act for reasons and engage in "practical" reasoning about what to do. Holmes might weigh the options of revealing his suspicions to the police or pursuing the one-legged man himself and decide to involve the police for the reason that they have the fastest boat. If action is the analogue to belief, we might similarly define it in part as a kind of response to our practical reasons, or as the conclusion of practical reasoning. From this perspective, action theory can be thought of as a counterpart to epistemology.

These are two somewhat different ways of situating action theory with respect to other areas of philosophy. We can motivate the inquiry even further by thinking about the practical relevance of the kinds of questions it promises

to answer. Often, it is of great importance to us to know whether something that happened was an intentional action or not. In the *Palsgraf* case, much depended on whether the fireworks were set off intentionally or merely by accident. Other related distinctions come up frequently in legal settings. The criminal law generally requires that there be a "voluntary act" in order for a crime to have occurred, for example. Merely planning to rob a bank is usually not enough to incur criminal liability. In addition, it distinguishes between acts that are undertaken with intent, with knowledge, out of recklessness, and out of negligence. It treats certain crimes that are committed as a result of passion or provocation differently than those committed with premeditation, but it treats those done in ignorance of the law and those done with knowledge of their illegality as the same. These all seem, at least in part, to be an effort to distinguish between different kinds of action. Philosophers of action aim to determine whether such distinctions hold up under scrutiny, and to articulate the precise conditions under which they apply.

There is a rich history of philosophical investigation into agency and action, reaching back to the very beginning of philosophy as a discipline. However, action theory was not really considered an autonomous area of philosophical inquiry as distinguished from ethics, free will, and philosophy of mind until relatively recently. The groundbreaking work of Anscombe and Donald Davidson in the mid-to-late twentieth century defined the set of questions that frame contemporary research on agency. This book aims to introduce readers to these contemporary debates rather than to offer a detailed historical perspective.

In spite of being a fairly young field – or perhaps because of it – there is quite a lot of disagreement about where exactly the philosophy of action should start and where it aims to end up. The organization of the material here reflects my perspective on the best way to progress through the topic, but this is one perspective among many legitimate ones. Nearly all of the chapters have a serious claim to being the one that should go first, and for each of those, there are respectable philosophers who would argue that it should be left out altogether. The goal of Chapter 2 is to lay out the thicket of overlapping questions, concepts, and phenomena that have been used to frame the investigation of agency. Close attention to these different starting points will equip the reader to better understand what follows, since in my view, many apparently substantive disagreements can be traced back to divergences in this first step.

Chapter 3 takes up the idea that the kind of action we should be interested in – rational or intentional action – is subject to a distinctive kind of *explanation*. When we ask why a person acted as he did, the answer can take the form of providing the person's reasons for so acting. Suppose we inquire into why Richie went to D.C. this weekend. If the explanation is that he boarded the wrong train, thinking it was going to Philadelphia, this reveals to us that he did not go to D.C. intentionally. "Boarding the wrong train" is not a reason to go to D.C. But if the explanation is that he went to see a concert,

the implication is that he did go there intentionally. One of the central debates in the field concerns the nature of this kind of rationalizing explanation. How exactly does citing a person's reason provide illumination of her action?

Chapter 4 turns to the question of what an action *is*. We learn from Chapter 3 that actions are the kind of thing that can be given a rationalizing explanation, but what kind of thing is that? This chapter delves into the metaphysical details. How are intentional actions related to unintentional ones? How are actions individuated from one another? Are some actions "basic," and if so, which ones? And most importantly, what is it that makes an event an action?

It is almost irresistible to think that the answer to this last question has something to do with what the agent of the action had in mind. Chapter 5 therefore shifts the focus to the notion of intention. We commonly speak of agents as *having* intentions or *acting with* certain intentions. What do we mean by these locutions? We might choose not to take such ways of speaking literally, viewing them instead as mere figures of speech. But those who do take them literally generally understand the term "intention" to refer to a kind of mental state or attitude. The challenge, then, is to understand what kind of attitude it is and how such attitudes are related to intentional action.

In a sense, Chapter 6 starts our discussion over. Like Chapter 3, it explores a phenomenon that many philosophers take to be a criterion of the kind of action we should be interested in: "practical knowledge." When we act intentionally, we normally know what we are doing in a distinctively first-personal way. For you to know that I am currently writing a book about agency, you will have to use methods like perception, inference, or testimony. You might lean over to take a glance at my computer screen or ask my colleague what I am up to. But I can know that I am writing a book without waiting to observe what appears on the screen or performing any kind of conscious inference. Reflecting on the phenomenon of practical knowledge gives us the opportunity to step back and further assess some of the proposals we have seen so far. For instance, some theories of intention seem to offer better explanations of practical knowledge than others. What we make of this depends on how central to our topic we take practical knowledge to be.

According to one kind of view, practical knowledge is not only central to our agency but the goal or "constitutive aim" of acting. The proposal is that in addition to the various specific purposes we have for doing things, intentional action itself has a purpose, namely, to afford us knowledge of ourselves. Other approaches agree that action has a constitutive aim but deny that this aim is epistemic. One longstanding tradition holds that all action aims at doing what is best, or at least what is good. Rival accounts claim that it aims at constituting or unifying the self, or at expressing the agent's drive to overcome resistance and expand her power. Against all of this, there are many who deny that action as such has any particular aim. Chapter 7 is dedicated to examining these debates.

Chapter 8 draws a distinction between mere intentional action and action "par excellence." The thought is that of all the things we go around doing intentionally or voluntarily, only some of them are truly autonomous. The

latter actions represent what the agent truly wants, or who she truly is, or what she is fundamentally committed to. To make good on this idea, the challenge is to specify which elements or structures of the agent's psychology truly speak for her, such that the actions that are suitably related to those elements count as fully autonomous. Is it some special subset of her desires? Her values or character? The plans and policies she has set for herself?

The counterpart to fully autonomous action is weak-willed or "akratic" action. We act weakly when we succumb to temptation or otherwise choose not to do what we believe would be best. We procrastinate, over-indulge, cheat on our vows, and work too much. The weak-willed agent is not ignorant of the considerations that speak against her action, and she does not act from vice – it is not that she takes the worse action to be the better. Nor is she straightforwardly compelled to act as she does, in the sense of having no choice in the matter. Rather, as we might say, she simply fails to control herself. But how is it even possible to act both intentionally and weakly, and what would it mean to exercise self-control? This is the central puzzle of Chapter 9.

Finally, Chapter 10 extends the focus to the case of acting together with other people. It is commonplace to speak of groups as deciding, intending, and doing various things. People play Scrabble, get married, elect presidents, and write laws together. Can we account for "shared" agency using only those resources mentioned in the first nine chapters? That is, can we understand it as a matter of individual agents with ordinary intentions who are each doing their part? Or do we need to introduce new notions like "group agents" or "group intentions?"

Though my own views on these issues will inevitably come through to some extent, I do not take myself to be arguing for any particular conclusions. The seminal texts are difficult enough that one cannot simply clarify what is going on in them without being somewhat opinionated. But given the divisions in the field, I will be satisfied if I can help to illuminate where different thinkers are disagreeing with one another and where they are talking past one another. My hope is that I have left sufficient room for readers to form their own views.

2 What Is the Problem of Action?

To begin our investigation into agency and action, we must first try to state more clearly and precisely what the central questions are. As we will see in subsequent chapters, there is a potentially bewildering array of different approaches to understanding what action is, or what it is to be an agent. This kind of vigorous debate tends to be the case in any area of philosophy, but the philosophy of action is particularly difficult in this regard. There is broad disagreement about the questions we should be interested in addressing and what our starting assumptions should be. To make matters worse, these disagreements are not always obvious or made explicit by the various parties in these debates, which can lead to theorists simply talking past one another.

The problem, I think, is that "action" is a sub-class of a more general category that we might call "activity" or "behavior." Any given philosophical investigation of action must begin with some idea of how this general category ought to be restricted. And there a variety of interests we might have in trying to draw these lines. Some think it obvious that our philosophical interest in action is broadly ethical, and so find ways of delimiting the topic that emphasize the connection to reason, responsibility, self-consciousness, and self-understanding. Others think the interesting puzzles are obviously metaphysical, and so frame the topic in a way that emphasizes the contrasts between actions and other kinds of events or occurrences in the world. Still others are interested in behaviors that are the manifestation of a certain kind of psychology, and so demarcate the topic in that way. The everyday notion of action is fluid enough to accommodate all of these purposes, and so we cannot simply rely on our intuitions about the meaning of the word to clarify what it is we are talking about. Yet for those of us who have spent significant time thinking about agency from a particular perspective, it is easy to forget that this is so (and I include myself in this).

This chapter will attempt to lay out the variety of ways that we might choose to frame the investigation, with the hope that we will be better able in subsequent chapters to understand why different theorists have ended up where they are.