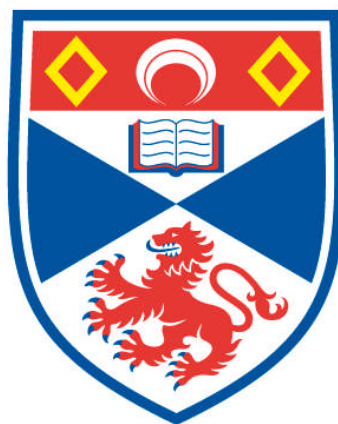


ACTING AND UNDERSTANDING

Alexander C. Blomberg Stathopoulos

**A Thesis Submitted for the Degree of PhD
at the
University of St Andrews**



2016

**Full metadata for this item is available in
Research@StAndrews:FullText
at:**

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/8417>

This item is protected by original copyright

Acting and Understanding

Alexander C. Blomberg Stathopoulos



University of
St Andrews

This thesis is submitted for the degree of PhD
at the University of St Andrews

September 2015

1. Candidate's declarations:

I, Alexander C. Blomberg Stathopoulos hereby certify that this thesis, which is approximately 6700 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2011 and as a candidate for the degree of PhD in the same year, the higher study for which this is a record was carried out in the University of St Andrews and the University of Stirling between 2011 and 2015.

Date signature of candidate

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date signature of supervisor

3. Permission for publication: (to be signed by both candidate and supervisor)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

No embargo on print copy

ELECTRONIC COPY

No embargo on electronic copy

Date signature of candidate

signature of supervisor

Abstract

This thesis concerns the question of what it is for a subject to act. It answers this question in three steps. The first step is taken by arguing that any satisfactory answer must build on the idea that an action is something predicable of the acting subject. The second step is taken by arguing in support of an answer which does build on this idea, and does so by introducing the idea that acting is doing something which is an exercise of a particular kind of disposition on the part of the acting subject. The third step is taken by arguing that the disposition in question must be of a kind which is exercised in conditions in which the acting subject thinks they are acting. From this vantage point the thesis develops many further commitments: That action is constitutively subject to a mode of explanation that mentions the kind of disposition just mentioned; that any case of acting requires a veridical representation of a means by which the action is performed; and that a problem about the underspecified nature of desire ascriptions can be solved by appeal to the conceptual materials made available by these investigations. The thesis finally develops several objections to the account it gives, both substantive and methodological, and explains why these objections ought to be rejected.

Acknowledgements

In writing this thesis, I was immensely helped by my supervisors. Alan Millar, though officially my second supervisor, has given very many grounded, lucid, and helpful responses to very much written output, especially early on. I thank Alan for helping me gain my footing. Adrian Haddock, my principal supervisor, has devoted very much time to discussing every topic of this thesis with me, and also to discussing many other topics. Adrian's way of taking my various claims very seriously, and his honed but vigorous and enthusiastic responses to these claims, have encouraged and helped me in finding the right way to express them. I thank Adrian for the resulting sense of confidence.

I also want to thank my friend Brian Ho, who has been very generous with his advice and variable-brow conversation, and other friends in Edinburgh, including David Silverman, Bren Markey, Bihotz Barrenechea, and Josh Thorpe. Then there are the other students, Zuzana Daňková, Pat McDevitt, Simon Payne, and Ryo Ito, and all philosophers at the University of Stirling, who made seminars and pubs challenging and fun.

I could not have completed this work without financial support from the University of Stirling, Erik and Gurli Hultengren's philosophy fund, Gunvor and Josef Anér's foundation, and the Royal Institute of Philosophy. I thank each institution for directing their generosity my way. I also thank those who have

read, listened to, asked questions about, or given advice concerning my written or spoken output, including especially Martin Gustafsson, who has encouraged me in the past and challenged some of my more recent pronouncements. And I want to thank those who have persevered in interacting with me through my distracted phases, including friends (especially Micke), and extended family (especially Maria, Petter, Kathleen, and Lilly).

Finally, because thanks would be inadequate and apologies inappropriate, I want to dedicate my work to my close family: Anna, Angelos, Amanda and Angelica; and to my partner Molly.

Contents

Introduction	1
1 Acting in Doing Something	8
1.1 Introduction	8
1.2 Doing Things	9
1.3 Acting in Doing Things	18
1.3.1 Acting Without Subjects	19
1.3.2 Acting Without Agents	24
1.3.3 Agents Doing Actions	29
1.4 Two Approaches	31
2 Doing What's Wanted	34
2.1 Introduction	34
2.2 Actions as Causal Relata	35
2.2.1 Psychology	36
2.2.2 Causation	37
2.2.3 Correspondence	38
2.2.4 Hopeless Deviance Problem	41
2.3 Acting as Exercising a Disposition	53
2.3.1 Exercising a Disposition	57

2.3.2	Exercising a Kind of Disposition	59
3	Doing What's Believed	67
3.1	Introduction	67
3.2	Practical Belief	68
3.3	Practical Belief is Imperfective	73
3.4	Practical Belief is Causal	87
3.5	Practical Belief is Part-Homogeneous	93
3.6	Practical Belief is Instrumental	96
4	Consequences	102
4.1	Introduction	102
4.2	Basic Action	103
4.3	Action Explanation	115
4.3.1	Constitutive Explanations	115
4.3.2	Dispositional Explanations	119
4.3.3	Nagellian Objections	123
4.4	Desire Ascription	137
4.4.1	An Equivocal Puzzle	137
4.4.2	On Desire Ascription	138
4.4.3	Semantic Revisionism	143
4.4.4	Metaphysical Revisionism	153
4.4.5	Deepening the Ordinary View	155
5	Circles	164
5.1	Introduction	164
5.2	Objections to Reflexive Belief	165
5.2.1	Chauvinism Charges	165
5.2.2	Possibility Protest	173

5.2.3	Objectivity Objection	175
5.3	Objection to Reflexive Accounts	176

Introduction

The question which much of this thesis is devoted to answering is what it is for a subject to act in doing something. Perhaps narrower versions of this question help spotlight its sense and import: What it is for someone to be *acting* in walking towards Rome, or in securing an all blue wardrobe, or (more pertinently to most) in destroying next year's crops through neglect? We can also make the question vivid through contrasting pairs of cases: Obviously there is a difference, which involves agency, between a stone's rolling down a hill and someone's picking it up — what is it? Or, to state the contrast in a different way: What is the difference between someone who is acting in dropping a stone and someone who is doing it unintentionally?

This thesis contains an answer to the overarching action-theoretic question, with these main three conceptual ingredients: Subjects doing things, their dispositions to do them, and (crucially) practical beliefs about things being done. The core of the answer may be introduced through a short series of questions (“Q”'s) and answers (“A”'s):

Q1 What is it for a subject, N, to act in Aing?

A1 It is for N *to A* because N *wants to A*.

Q2 But what is it for N to A because N wants to A?

A2 It is for N to exercise a particular kind of disposition.

Q3 But which kind of disposition is it that, when exercised, amounts to an agent acting in Aing?

A3 It is the kind of disposition such that, in a particular kind of condition, this subject acts.

Q4 But now what is the condition by which we are to distinguish this kind of disposition?

A4 The presence of a belief which represents this subject as acting in Aing.

So my answer begins as a platitudinous form of psychologism, but then takes on the shape of a minority dispositionalism, which then incorporates an insight about the role of self-awareness in action.

For A1 seems to be the closest thing we have to an action-theoretic platitude. It is, it seems, a commitment of what gets called the “standard” account of action, on which someone acts just in case there’s a desire and a belief, and perhaps some other bit of psychology, which jointly cause an event, and, as will probably be added, “in a specific way” (Velleman, 1992; Smith, 1994).¹ But A1 is more widely acknowledged than this kind of account, which is really just a particular way of developing it.²

Adding A2 moves us beyond this platitude to a somewhat neglected minority dispositionalism. For a smallish number of contemporary thinkers, like Davidson (2004, p. 108), Hornsby (2008, p. 5), Coope (2007), and Hyman (2014), have seemingly attempted to give sense to the platitudinous claim by construing

¹I mean that “wants to A” can be taken as rough talk for whatever specific sets of attitudes such theorists imagine.

²A1 is also seemingly earlier acknowledged than the standard account. As far as I can see, contemporary figures have been able to derive support for A1-conforming theses from such diverse historical figures as Aristotle (Coope, 2007), Aquinas (Anscombe, 1963) Kant (Rödl, 2007) and Hume (Smith, 1994), and so all of these authors can, I think, be seen as precursors to the vague “psychologism” which A1 expresses, though perhaps only Hume, or Hume and Aristotle, may be seen as anticipating the standard account.

desire as a disposition, power or capacity, exercising which amounts to acting. These theorists do however tend to be silent on the nature of the kind of disposition they thereby introduce, and accordingly my agreement with them only extends so far.

Adding A3 and A4 articulates what it is to have and to exercise that kind of disposition. It does so by pairing two thoughts. The first thought is that a disposition is to be distinguished through mention of the kind of condition in which it manifests and the kind of doing which counts as its manifestation. Hence a desire must be a disposition such that, in some condition, its subject acts. The second thought, closely associated with Anscombe (1963), often rejected by “causalists”, and seldom considered by “dispositionalists”, is that it is a mark of the fact that someone is acting that they believe it. The thoughts are paired by making such Anscombian awareness not merely a necessary condition for someone’s acting, but the condition in which a subject who wants to act does so. Hence a desire is viewed as a disposition such that, when a subject believes they are acting, they are acting.

Many questions will arise about this answer to the action-theoretic question, not least of which are those which probe the circularity of the account. In the thesis I develop some such concerns briefly and some at length, and try to give satisfactory responses to them. I also hope to show that the present account helpfully contributes to a few difficult issues, which concern, for example, basic action, action explanation, and desire ascription. I describe these objections and virtues in the following summary of chapters:

The *first chapter* of the thesis is devoted to (what should be) the platitudinous claim that, for N to act in Aing involves, among other things, for N to be Aing. It provides an account of what it is for N to be Aing on which it is for a certain kind of predicable to attach to N. It then looks at accounts which try to

understand what it is for N to act in Aing by introducing a doing on part of an indeterminate subject, or by introducing a doing on part of N's body. It argues that these accounts fail. Predicating an action of an indeterminate subject, or of a subject different from N, cannot do justice to the idea that N is doing the action. But a different and more radical form of account remains, which does not view an event of Aing as something that is predicable of N, but rather as something related to N through some *two-place* predicable.

The *second chapter* starts by engaging with accounts which are all, it seems, complicated versions of that relationalist form of account. These accounts try to understand what it is for N to act in Aing by supposing psychological relations (that N has various psychological states which *represent* Aing) and a causal relation (that these states *cause* Aing). But so-called deviant causal chain cases show that these conditions are not enough, and I argue that, in trying to add extra conditions to complete this form of account, many theorists fall prey to more or less complicated versions of a fairly simple dilemma: Either the extra condition is tacitly introduced *merely as* a condition which defines acting, making these accounts uninformative, or the extra condition is given its own sharp edges, but thereby fails to provide the right definition of what it is to act. The chapter goes on to outline a different kind of account, on which for N to act in Aing is for N to exercise a particular kind of disposition — properly called a desire —, individuable as the disposition which is exercised in a particular kind of condition. The chapter ends by suggesting that the condition which individuates this kind of disposition is the presence of a belief which represents its subject as acting.

The *third chapter* starts with the working hypothesis that an account could be given which said that acting in Aing is exercising a desire to A, where a desire to A is a disposition to do what satisfies a practical belief in the presence

of such a belief. It then argues that if a practical belief is construed as a belief that its subject is *acting in Aing*, this completes an account of that form. Most of the rest of the chapter is devoted to unpacking the content which a practical belief needs to be assigned if the thesis that associates a practical belief with every action is to survive scrutiny. It emerges that a practical belief which represents its subject as acting in Aing needs also to represent this subject as doing something else which causes Aing, and as acting in doing that other thing, and as doing that other thing *in order to A*.

The *fourth chapter* is devoted to extracting three consequences from the account, which are treated in three respective sections. The first of these sections describes a consequence for the metaphysics of action: The account makes it impossible to do an action, without doing another action in order to do it, in the belief that one is. I consider and reject some arguments which are supposed to generate a need for introducing actions which nothing is done in order to do, as well as some arguments which are supposed to generate a need for introducing actions which are performed without believing another action is done in order to do it. The second of the sections concerns a set of issues in the explanation of action. One, raised by Nagel, is of roughly this shape: What sort of insight is provided by saying an action is done on account of being desired, if its being done on account of being desired is entailed by its being an action? I argue that we should learn to live with the idea that “N wants to A” lacks explanatory import to someone who knows that N acts in Aing. But I argue that mentioning *larger* bits of someone’s set of desires (“N wanted to A in order to B”) can provide insight as to *why* someone did B or *how* they wanted to A. This thought ties in with a discussion in the third section, which concerns a consequence about attitude ascription: It is commonly supposed that the objects of our desires are normally revealed right on the surface of our ascriptions, so that “N wants

to go fishing” ascribes a desire with the simple object, and hence the simple satisfaction condition, that N ends up going fishing. The section defends this common supposition against some recent attacks, which turn on the idea that desire ascriptions are often in some sense incomplete, revealing only part of what the agent really wants to do. The section shows a way to reconcile the commonly supposed view with this idea, appealing to the idea that a desire ascription (like “N wants to A”) sometimes specifies just a part of an instrumental chain of desires (such as that N wants to A in order to B).

The *fifth chapter* is devoted to defending the invocation of a belief about acting in this account of acting. I develop and undermine some complaints, which rest, for example, on the idea that it is somehow empirically precarious to suppose that every acting subject must believe that they are acting, or on the idea that this generates infinitely recursive beliefs which must, because of their infinity, somehow be out of reach to an ordinary subject. The only really good objection, I hold, is the one which questions the *informativeness* of the form of account which says that for N to act in Aing requires N to believe this very thing. My response to this serious worry involves contrasting the form of account which simply says that for N to act in Aing requires a psychological state which represents N as acting in Aing (proposed by such theorists as Harman (1976); Searle (1983); Setiya (2007), and often left unchallenged), with my preferred form of account, which adds to this that for N to act in Aing requires psychological states which represent N as acting in doing *other* things envisaged to have suitable instrumental connections to Aing. The latter account is better than the former in roughly the same way that it is better to say that “believing that p is believing other things which cohere with p” than it is to say “believing that p is holding p to be true, where ‘holding true’ means believing”. One difference between the two forms of account is that the former, worse kind of

account allows an interlocutor to keep asking one and the same question “what is this concept of ‘acting in Aing’ which figures in the thought of someone who is doing that?”, whereas the latter, better kind of account does not allow an interlocutor to keep asking any single question. And if the account given allows no single question of the form “what is acting in ...ing?” to be raised more than once, it does not leave any single question of that form unanswered.

Chapter 1

Acting in Doing Something

1.1 Introduction

This chapter formulates the action-theoretic question which the two subsequent chapters are devoted to answering. It formulates this question as follows: What is it for some subject to be acting in doing something? It offers a determinate conception of what it is for a subject to be doing something, on which doings are a special sort of predicables. It then argues that two types of account of the nature of action fail because they do not do justice to the idea that actions are predicable of the subjects doing the actions. The first kind of account builds on the idea of a deed with an indeterminate subject. It must fail to do justice to the idea that agents are *doers* of their actions. The second kind of account builds on the idea of a deed of a subject, but not the subject doing the action, instead casting the subject's body in a primary role. Again, such an account must fail to do justice to the idea that *agents* are the doers of their actions. Finally it is suggested that a more radical form of account remains, which simply does not accept my way of formulating the problem — does not view the action-theoretic

question as a question of what it is for a subject to act in doing something. This alternative account is likely to proceed, instead, to try to account for relations which it assumes makes a bit of behaviour, or an event, into someone's action.

1.2 Doing Things

Action theorists commonly proceed by assuming some capacious notion of happening within which they seek to delimit action. For example, Davidson worked with an "ontology of events" and wondered, at times, how an event need be related to a subject to be that subject's action (Davidson, 2001a, p. 43). In sympathy with Davidson, though not always in explicit agreement with his doctrines about events, many have started with the apparently narrower notion of a "bodily movement" and asked what makes some bodily movement into someone's action (Smith, 2012, pp. 387-389). Some have started, instead, with the different, very broad, but basically unelucidated notion of "doing something", and asked what makes it the case that a subject is doing something *intentionally* (Setiya, 2007, pp. 23-24).

Aligning with this practice, and particularly with Setiya's example, I will seek to circumscribe a notion of *acting* within a more capacious notion of *doing*. But I hope to take less for granted than is usually done about the nature of the capacious notion of my choice. Hence in this section I'll say something, hopefully uncontroversial, yet informative enough to delineate what it is to be doing something. In the next section, I pose the question of what it is to be acting in doing something.

On the face of things, at least, we can circumscribe the topic of doing something through the following distinction: There is, on the one hand, how things are, and on the other, what things are doing. How things are includes such things as that something is red, or that someone is at the top of a hill. What things

are doing includes such things as that something is turning red, or that a spider is making a web. The applicability of this distinction presupposes that being in some way and doing something are alike in so far as both are predicables.¹ But if the distinction is marking anything more than a verbal difference, it must be possible to make some sense of the suggestion that these predicables are of distinct kinds. In this section I want to take seriously both these suppositions: Doings are predicables, and they are predicables of a distinct kind.

We want an answer to the question “what is it for something to be doing something?”, as I suppose we would like an answer to the question “what is it for something to be in some way?”. But we should be careful which type of answer we wish for. It seems incredible that there should be a reductive answer to the question about being, and I think we should be similarly incredulous in the case of doing.² Given that, what can be said about what it is to be doing

¹In OED-influenced language, we may say that a predicable is something that we may truly or falsely affirm or deny of a subject. This definition conveniently obscures whether predicables are linguistic predicative phrases used to say that individuals are such-and-such, or concepts through which we may think of individuals as being such-and-such, or ways in which individuals can be such-and-such. It is tempting to say our topic is metaphysical, not linguistic or conceptual, since it is tempting to say our topic is running, biting, nailing and sinking, etc., and not so much “running” and “biting”, etc., or concepts of running or biting. But this is somewhat premature, since one can *define* a type of linguistic predicative phrase as one which concerns things that things are *doing* (which seems to me to be pretty much how we learn to use the phrase “verb phrase” — except that stative expressions like “loves Adam”, “is green”, “equals five” and “believes that everyone is equal” are then shoehorned into that linguistic category on grounds of surface similarity), or *define* a type of concept as one which concerns things that things are doing (and I have no hold of how to define a type of concept except as concerning some type of situation). Given these possibilities of aligning linguistic or conceptual distinctions with metaphysical ones, there does not seem to be anything necessarily wrong with conducting the investigation in an overtly linguistic mode, or by talking very much about “concepts”. Such ways of proceeding, though sometimes unattractively indirect, seem at points to be the only ways to avoid extremely cumbersome phrasing. For example, when below I speak of when it is licensed to “move from” claims about events to claims about doings, I do not exactly mean when that would be sanctioned by English as it happens to be spoken, or even as it should be spoken, but that it would be sanctioned by the kind of metaphysically felicitous English I am trying to speak. It would seem somewhat unattractive, and would probably invite some blank stares, if I always opted for phrases like “for someone to be doing something is for that something to be happening”, though I do sometimes opt for phrases which are only slightly less convoluted.

²For an example of a claim which might be interpreted as a reductive account of the former, consider logical atomism’s idea that states of affairs are somehow “made of” “simple” “objects” (Wittgenstein, 1922, §2.0272). For an example of a somewhat analogous claim about doing, though it probably isn’t intended to be reductive, consider Hornsby’s (2012) idea that doings are somehow “made of” a “stretch” of a given type of “activity”. If construed as reductively

something? I think we'll do best by taking for granted that what things are doing are predicable on a par with how things are, and exposing a feature that encompasses and is peculiar to what's in the first category. Using "N" as a stand-in for any subject, and "Aing" as a stand-in for any predicable of doing, I think we have such a feature in the following:³

Doing: For N to be Aing is for N to be Aing for a stretch of time, after which N will either have A-ed, or failed to

This feature is most clearly exemplified in cases like the following, which I take to be paradigmatic: Something is sinking to the bottom of the ocean, or someone is making an omelette. If a subject is doing either of these, this subject is bound to be doing so for a time, and once such a subject has stopped doing either, it is bound to either have succeeded in doing it (the thing having sunk to the bottom of the ocean, or the cook having made the omelette), or failed to (as, for example, if the sinking thing got caught in the fin of a dolphin, so that it never sank, or if the omelette was scrambled, so that it was never made).

Not all cases are like the paradigmatic ones above. But on closely considering these other cases, we'll discover that they too yield to *Doing*. A large class of the cases which need to be considered are those commonly called "activities", which include things like walking, skiing, and deteriorating. Of these doings, it has been observed that we can never truly say things like "this person was walking, but never walked", or "their skills kept deteriorating, but never deteriorated".

And this seems to show that a subject cannot end up failing to have done one

explaining what it is for something to be in some way or doing something, these accounts will be simply baffling. No amount of harping on sets or arrangements of primitively understood objects seems to tell us what it is for something to be red, just as no amount of thought about quantities of primitively understood "event goo", will seem to tell us what it is for something to be rolling down a hill.

³The rest of this section owes a lot to the relatively small number of philosophers who have made distinct contributions to action theory by reflecting unusually closely on the notion of doing something, with a special focus on the imperfective aspect, including Mourelatos (1978); Falvey (2000); Rödl (2007); Thompson (2008); Hornsby (2012).

of these things. But as soon as we see this, we also see that the subject doing an activity must end up having succeeded in doing it. For it seems we *can* always say “this person was walking, and ended up having walked”, and the like. Hence activities don’t, after all, evade our thesis about doing. And now it seems that we can easily characterise activities in a way which makes explicit how they conform to our thesis about doing: If a subject is Aing, where this is an activity, then this subject has already A-ed.

There seems to be a different kind of doing, commonly called “achievement”, which shares the above feature of guaranteed success, but for a different reason. For we can never say that someone is, say, hitting the ground, or outliving their arch enemy, and yet doesn’t end up having hit the ground or having outlived this enemy. But the explanation we just employed for activities is not available here, as it hardly follows from the fact that someone is outliving their arch enemy that they *have* outlived this enemy. The explanation which brings achievements under the heading of doings, while preserving their feature of guaranteed success, is not far off, however. For though outliving someone doesn’t require having outlived them, it does seem to require future success in so doing. And this seems to be a general feature of achievements. So we may say that if a subject is Aing, where that is an achievement, then although this subject hasn’t A-ed, it *will have* A-ed.⁴

⁴Maybe we could say that someone is performing the literally Sisyphean task of rolling a stone up a hill, though, of course, if it is literally Sisyphean, they could never end up having rolled it up that hill. Like activities and achievements, such doings still conform to the claim about doing, since after the time during which someone is doing such an impossible thing, they will have failed to do it. A possibly different type of case, discussed by Rödl, is that of staying healthy. Rödl says this is an “infinite end” in the sense that when it comes to staying healthy, “the contrast between pursuing and having got does not apply” (2007, p. 36). Perhaps this means that if one is staying healthy, one has already stayed healthy, and then I suppose the case would fit in with activities like running, walking or skiing. A different way of understanding staying healthy might be by saying that the idea of succeeding in doing it finds no real application after the time during which one has been keeping it. But then it seems possible to understand staying healthy as on a par with the literally Sisyphean task of rolling a stone up a hill — a doing that is guaranteed to end up in failure, but has not yet done so.

Some have proposed to explain the nature of achievements by saying that they take no time, or (if that is different) by saying that, when it comes to those doings which are achievements, the question of how long they go on finds no application. Ryle takes the latter stance:

We can ask how long it took to run a race, but not how long it took to win it. Up to a certain moment the race was still in progress; from that moment the race was over and someone was the victor. But it was not a long or short moment. (Ryle, 1949, p. 302)

If the Rylean account of achievements is true, then it won't be possible to bring these under the scope of *Doing*, since that thesis requires that a doing has a durative character, whereas Ryle seems to say that achievements lack durative character. One possible way of for me to accommodate achievements, as Ryle construes them, would be to change the letter of *Doing* so that it did not require all doings to go on for a time, but only that they go on at a time. This would accommodate the Rylean account, and preserve much of the spirit of *Doing*, as I do not think that much of what I'll say in this chapter hinges on whether achievements are taken to take time or not. (The crucial distinguisher is the idea of success and failure.) But Ryle's account seems implausible to me, whereas *Doing* seems in order. Below I say why I want to persevere with a durative conception of achievements:

There are two interpretations of the Rylean account of achievements. On one interpretation, no one is ever *doing* an achievement, although lots of things have done achievements. On another interpretation, something can be *doing* an achievement, but not during any particular span of time, so that perhaps achievements are temporally point-like. I am not sure that there is a sensible and coherent way to make sense of such non-progressive or non-durative doings. But even if there is, it simply seems correct to me that someone can be winning

a race during a time up until the point where they have won it, or that someone can be outliving their enemy during a time which lasts until they have outlived them. If so, Ryle's account, on either interpretation, must be false. Mourelatos (1978, p. 417) seems to share my view, and supports it by citing linguistic evidence. But to me such evidence seems less secure than the view it is meant to support; Mourelatos's view reads like common sense to me, and does not seem problematic in any clear sense.

I have hinted that activities and achievements can be understood as somehow derivative of paradigmatic doings. Defending that claim is not strictly necessary for defending the above thesis about doing. But the claim would help undergird the thought that all doings, including these ones, make up a smoothly unified category. So here is a development of the hint: When a subject's paradigmatic doing (such as that N is *walking to school*) is *presented* with the presumption of ultimate success ("N is *arriving* at school — any hour now!"), then it is *presented* as an achievement. When a subject's paradigmatic doing is *presented* as to be already partially successful (in a limited sense which sets the bar for success so low that it is guaranteed by what is currently happening) it is presented as an activity. ("N is *walking towards school*, and has *walked towards* it for quite a while now!"). So activities and achievements are not so much special doings that cannot fail, but paradigmatic doings presented with a view to future or past successes. Attributing achievements is attributing paradigmatic doings with a presumption of future success, attributing activities is attributing paradigmatic doings with a view to recent partial success.⁵

⁵This view does not *clearly* address, for example, cases where something is standing still. For where is the paradigmatic doing, capable of both succeeding and failing, in a case where someone is standing still? Perhaps it is a good enough answer that the subject is standing still for a time that extends to some indeterminate point in the future, which will either be reached while still standing still or not, so that that episode of standing will either succeed or not. This answer makes it a fuzzy question, perhaps, in what conditions this doing succeeds or fails. But it defends the idea that someone who is standing still counts as doing that only because there is something — standing still for 'a while' — which they will have succeeded or failed to do in the future. Admittedly some might feel a need to question this method for

The above defends the thought that *Doing* encompasses all doings. Now it is time to defend the thought that *Doing* describes something that is peculiar to doings. This is quite straightforward. It is not so much the bit about durations which sets predicables of doing apart from those of being. For putting aside frail specimens like “green at this point-like moment in time” or “*presently* winning”, every predicable which applies to a thing *does* seem to apply to it during some stretch of time. But the notion of progress leading to success or failure — Aing until having A-ed or failed to — finds no application in cases where a thing is blue, dead, sad, knock-kneed, or where someone believes Jordan is on Jupiter. If something is, say, blue, but then stops being blue, then it did not succeed in being of that colour, nor fail to. It just was blue, and then wasn't.

A hurdle, briefly hinted at in note 1, is that several psychologising phrases, like “believes ...”, “loves ...”, etc., have the look and feel of verb phrases, but do not seem to report doings on a par with the others we haven't noted. This presents a difficulty to those who want to make their distinctions from a linguistic point of view, since much of what we want to say about other verb phrases does not apply to these.⁶ But since we have not defined what it is to be doing something as being the subject of a verb phrase, distinguished by some syntactic test, we have the advantage of not needing to say that believing, loving and the rest are things subjects can be doing, in our preferred sense.

Nevertheless, it might be thought that something needs to be said about the surface similarity between “N is running to school” and “N believes the school is near”, and the corresponding surface difference between these two and “N is red”, which does not assimilate believings and lovings to doings. I think all that is needed is to emphasise the relative unimportance of grammar. It certainly is

accommodating the case. But let's leave this potentially complicated matter at this, since, again, the above thesis about doing requires only that all doings, including activities and achievements, involve doing something for a time after which there's either been success or failure in doing it.

⁶See Mourelatos (1978) in this connection.

possible, and might have been more frequent, for the language of doing to be adopted to report facts of being, and vice versa. What we express by saying “N is running to school”, we might, if convenience allowed it, have expressed by saying “N is presently a to-school runner”. What we tend to express by saying “N is red” we might have expressed in a Heidegger-evoking verbalisation, such as “N is redding”. It seems to me that “N believes Jordan is on Jupiter” is similar to such potentially misleading uses of language. The expression adopts a form which suggests that N is doing something to report some very different sort of fact about N.

The trouble is that we are not sure what we should say it does report. Is it that N is in some way, as when we say N is red? (Is believing being in a state?) Or should we say that knowing, loving and understanding belong in a category all of their own, beside the categories of doing and being? To these questions I have no answers, except the feeble answer offered above: Psychological predicables are unlike doings and like ways of being in the specific way that ongoing engagement leading to success or failure finds no application with them.⁷ (In the following chapters I will sometimes speak of psychological states, but only because “psychological predicables” sounds particularly linguistic, and is anyway quite non-standard. I do not think anything in my discussion will hinge on using this form of expression.)

It should then be noted that just as, in general, there is such a thing as *bringing about* or *maintaining* the having of a property (keeping oneself rich, of the colour green, or unwounded), there might be such a thing as bringing about or maintaining these psychological predicables (maintaining love, or bringing about belief in God). But making oneself believe in God, if that is possible, is something different from believing in God, and only the former can properly

⁷See Vendler (1957); Mourelatos (1978) for more detailed expressions of uncertainty concerning psychological verbs.

be called something which one is doing. All of this is of some relevance to the question of whether and in what sense it is impossible to “believe at will” (Hieronymi, 2009; Setiya, 2008a). For it gives us a simple argument for this attractive but elusive thesis, and assigns it a modicum of clear sense: If acting is doing something, and believing isn’t doing something, no one can act in believing — though it is left open whether someone can act in *maintaining* a belief.

I’ve said that for something to be doing something is for a special kind of predicable to attach to that thing. I want to proceed from that thought to pose the question of what it is to be doing something that is an action, or to be acting in doing something. As I briefly hinted earlier, many who are working in action theory prefer to speak not of things doing things, but of events (or “activities”, “doings”, “behaviours” — the important shared feature is the seeming quantification over items). These theorists try to proceed from there to ask questions of the sort: “What does it take for an event to be someone’s action?” As a gesture of pre-emptive diplomacy, I want to note that it seems possible to introduce a notion of event which is entirely parasitic on the notion that something is doing something, and thereby reconciles these questions. For we can say that what it is for there to *be an event of Aing*, as such theorists are liable to say, is for *something or other to be Aing*. Put in other words, thinking of an event just is thinking of what something is doing in a way which abstracts from the doer. Assuming this notion of an event, we can move frictionlessly from claims about what things are doing, like “the toothless wolf is eating the last grape”, to claims about events taking place, like “the last grape is being eaten”, and from claims like “there is an event of the last grape being eaten” to ones like “something is eating the last grape”.

The assimilation is helpful because it allows us to provisionally translate the

work of the kind of action theorist who speaks merely of events, into our adopted language of subjects doing things. Where they say “there’s an event of ...” we may say “something is doing ...”. To boot, it allows us to translate their apparent talk of *types* of event, like “movement” or “oil-painting”, into talk of *types* of doing, like “something’s moving” or “someone’s painting in oil”, and finally their talk of events which — in a sense *they* often don’t try to articulate — are associated with a particular type of material substrate, like “*bodily* movement” or “*fungal* spore spreading” into talk of movement done by subjects which are bodies, or spore spreading done by subjects which are fungi.

1.3 Acting in Doing Things

If something, “N”, is doing something, “A”, then there are these two possibilities: N is *acting in Aing* (where as far as I understand them, this phrase is getting at the same thing as Davidson’s (2001a) “Aing *intentionally*”, Setiya’s (2007) “*doing something intentionally*”, and Velleman’s (1992) talk of “when someone acts”). Or N is merely Aing, and not acting in Aing. The question that will occupy us for a while is the nature of the first possibility: What is it for N to be acting in Aing?⁸ Given what was just said about viewing events as doings, I think this question should be seen as a clearer way of articulating commonly

⁸A note on the use of variables: This question is clearly not asking about some particular subject what it is for it to act in doing things, or about some particular type of thing that can be done when it constitutes acting. Rather, “N” is used as a stand-in for any subject, and “A” for anything that N is doing. So the question requires an answer in the form of a general specification of what it takes for the world to be one in which *a given* subject is acting in doing *a given* thing which this subject is doing.

If there is a general answer like that, it will, of course, answer all specific questions which can be generated by replacing the variables with a specific subject and doing, such as “what is it for the gardener to be acting in misreporting their work time?”. It will also answer intermediate questions which can be generated by just narrowing the variables, as in “what is it for a non-human animal to be acting in moving its limbs?”. Questions about specific things, or ranges of things, acting in doing specific things, or ranges of things, will then come out seeming directly dependent on the main action-theoretic question posed. To answer them we’ll need to know whether the specific subject or kind of subject mentioned is standing to the doing or kind of doing mentioned in the way that subjects stand to doings when they are their actions.

posed questions like “which events are actions?” or “what is it for something that happens to be some subject’s action?” — clearer, at least, than these questions are when posed in the absence of any determinate conception of event. Building up to a full answer to the above question, I want to defend a partial and very platitudinous-sounding answer, missing which seems to lead a lot of people astray. The partial answer is that for N to be acting in Aing is, among other things, for *N to be Aing*.

1.3.1 Acting Without Subjects

Many theorists give theories of acting in which they speak of “events” without introducing subjects as the doers of these events. If I understand these accounts, they say that for N to be acting in Aing is a matter, at least, of some event, A, happening. But they do not say that it is N that does it.

It is worth noting that such a theory could be defended in one very non-committal way and at least one very committal way. The basic and non-committal idea just is to *avoid* speaking of *N as Aing* in the account of what it is for *N to be acting in Aing*, adopting, instead, Cluedo-esque talk of “the Aing that occurred or is occurring” or “that episode of Aing”. Thus far, someone proposing this approach *might* still agree with me that things that are happening are things subjects are doing, and accordingly it would be an *add-on* to such a theory, and would make it much more committal, to *deny* this, and opt, instead, to introduce a different conception of what it is for something to be happening — perhaps by introducing an ontology of individual, unrepeatable, historical events in the manner of Davidson. Such further moves lead to restrictions and complications in how the action-theoretic question is asked and answered, some of which generate challenges to their proponents. But the possibility of a noncommittal view shows, I think, that it would be a mistake about scope to

argue against, for example, Davidson’s specific commitments about actions and events, thinking oneself thereby to undermine the whole idea of accounting for N’s acting in Aing without predicating Aing of N.

A different objection, due to Thomas Nagel, seems to be of that wider scope. For he quite directly says that an account which only mentions a subjectless “flux of events”, can’t provide for the idea that actions are “assignable to individual agents as sources” (Nagel, 1986, p. 110).⁹ Nagel embroiders his objection as follows:

[M]y doing of an act — or the doing of an act by someone else — seems to disappear when we think of the world objectively. There seems no room for agency in a world of neural impulses, chemical reactions, and bone and muscle movements. Even if we add sensations, perceptions, and feelings we don’t get action, or doing — there is only what happens. (Nagel, 1986, p. 111)

I think Nagel’s objection is correct on a natural if somewhat pared down reading. But Nagel does not explain why, and says things which might mislead us as to why.

Obviously the above quote is more an expression of disbelief in accounts of the above form than an argument against such accounts. Failing to see much of an argument in Nagel’s surrounding comments, others have confidently asserted that all that is missing from the targeted account is some mention of the (probably psychological) states of the agent implicated in the event which

⁹His suggestive but brief comments have been picked up by Hornsby (2004; 2008), Coope (2007) and Velleman (1992). Hornsby thinks the objection is fatal to a number of theories of what it is to act, whereas Velleman thinks it only encourages some additions. But as Hornsby (2008, pp. 7-8) brings out, Velleman only seems interested in the same problem as the others near the start of his article, where he presents the Nagel-evoking thesis that if we merely tell a story in which “an intention causes bodily movements” we will wrongly represent things as if “nobody — that is, no person — *does* anything” (Velleman, 1992, p. 461). Soon after this Velleman seems to slip into asking and answering what looks like a very different question, concerning “that which distinguishes human action from other animal behaviour” (Velleman, 1992, p. 462).

is the action. “Agents are not left out[...] when the relevant events consist of the agent’s being in certain states”, says Davis (2010, pp. 33-34). And Bishop concurs:

Accepting that intentional actions have active mental causes certainly dispels the charge [of] ‘leav[ing] the agent out of the picture’.
(Bishop, 2012, p. 61)

Davis and Bishop seem to suggest that there is no problem of disappearing agents once we supplement the kind of account Nagel targets so that it amounts to something like this: “for N to be acting in Aing is for Aing to be happening (where this doesn’t entail of anything in particular that it is Aing) due to the fact that N is in psychological state X”.

It seems obvious that Nagel would not be satisfied with this. If his problem was simply with the lack of mention of an agent in the targeted accounts, it would, of course, be solved once an agent was mentioned, but then what would be the point of raising it? He also clearly says that mention of psychological items can’t solve his problem. It would then be surprising if the problem was such as to be solved by combining mention of an agent with mention of some psychological states (fuzzily envisioned to “belong” to that agent). But all this does little to alleviate the fact that Nagel does not explain how his problem is distinct from the one his targets claim to have overcome.

A further obstacle to understanding Nagel is that his comments on the issue are liable to distract us from the problem and towards his diagnosis of it. Nagel thinks the problem is generated by a larger difficulty of his interest, which concerns how to combine “the perspective of a particular person inside the world with an objective view of the same world, the person and his viewpoint included” (Nagel, 1986, p. 3). More specifically, his diagnosis seems to be that, by taking an “objective” perspective, we can only discover a subjectless “flux of

events”, whereas agents who are acting in doing things can only be discovered through another, “subjective” perspective. The targets of Nagel’s criticism may be forgiven for trying to transform these unwieldy concerns into something they know how to respond to, though they should not get a pass on the fact that they thereby seem to solve a problem not worth raising.

We should hold out hope for a way of explaining Nagel’s problem without becoming engrossed with his diagnosis. For it seems quite clear that what Nagel says about perspectives does not elucidate the centre of the problem, which, to remind ourselves, is really just *why* a subjectless flux of events (perhaps circumscribed by saying it involves psychological states “owned” by a certain subject and standing in certain causal relations to certain events) couldn’t felicitously be identified with an agent’s participation in their action. The thoughts “if I view the world third-personally, I only discover a flux of events” and “if I view the world first-personally, I discover agents who are acting in doing things”, do little to defend or explain the thought that “a flux of events can’t constitute an agent who is acting in doing something”. We may be encouraged by this to cut past talk of perspectives: Rather than saying the problem concerns giving an account of “subjectively encountered facts about subjects who are acting” which mentions only things in “an objectively encountered reality of neural impulses, muscle movements, and other things that happen”, we may just say it concerns giving an account of facts about subjects who are acting in terms of subjectless facts about events. How should this simplified version of the problem be undergirded?

As a first step towards defending Nagel’s pessimism, suppose, as we seemingly can’t congruently deny, that for N to be *acting in Aing* requires N to be *Aing*. If this simple thought is true, then no account which fails to entail that N is *Aing* can be a true account of what it is for N to be acting in *Aing*. But now

one may wonder how an account the point of which it is to avoid saying that N is Aing, in favour of speaking subjectlessly of events, is supposed to defend this implication. The trouble is that there does not seem to be any entailment running from such facts as that N is in some psychological state which causes Aing to be happening, to such facts as that *N is Aing*

I earlier registered doubts that it would be possible to give a general reductive answer to the question of what it is for something to be doing something, but that is not quite at issue here. The present proposal is weaker in that it only requires one-way relations of entailment from the theorist's favoured causal-psychological facts to facts about the doings of subjects. But the apparent distinctness and fundamentality of the notion of doing things still makes it hard to overcome the feeling of a gulf. As Nagel and some of his followers — especially Hornsby — keep saying,¹⁰ a bunch of states and subjectlessly identified events, however large, and however causally related to each other, seem to be hopeless materials for building up to the idea that agents are the doers of their actions. The trouble is that Nagel and his followers are of little help in explaining why this need be hopeless, as most of what they say on the matter just repeats the basic conviction.

So what is the completing step to defending their pessimism? My argument is flat-footed, but hopefully therefore readily acceptable: Since, in general, talk of a subject, a feature, and some causal relationship between the subject and the feature, can't entail that the subject has the feature, mention of N, Aing, and a causal relationship between N and Aing can't secure that N is Aing. The general point is readily illustrated: "Here's N", "There's thinness", and "N brought the thinness about", hardly entails that N is thin, and adding features to N or complicating the causal relationship between N's features and the resultant thinness hardly improves the prospects of securing the entailment. (Note that

¹⁰Especially in Hornsby (2004, p. 12).

it doesn't matter to this flat-footed argument whether we think of thinness as an entity with a life of its own, and hence not as a predicable, or, instead, as a predicable, though without predicating it of anything determinate.) Why should things be thought to be any different when it comes to what subjects are doing?

1.3.2 Acting Without Agents

Other theorists give accounts of what it is for N to be acting in doing things in which, seemingly, they accept the idea that things do things, since they often refer to what N's body is doing, though they avoid mention of what N is doing. I don't know whether the preference for body talk is the *reason* for the neglect of agent talk, or if, rather, some doubts about the permissibility or usefulness of speaking of N as *doing things* in an account of what it is for N to be *acting in doing them* leaves these theorists with N's body as the only eligible-seeming subject on which to pin the things that, after all, seem to happen when subjects act.¹¹

This approach to action is not always explicitly stated, and some theorists seem to vacillate on the matter. But there are clear and prominent instances.

One is Danto's declaration that "an action [is] a movement of the body plus x,

¹¹I can think of other rationales for this kind of proposal, but none is very complete or very good. I offer some here: Perhaps the idea is that a focus on the activities of bodies makes room for a certain kind of conjunctive analysis, discussed by Ford (2011), which would otherwise seem impossible: Saying that N acts in Aing just in case N's body moves plus p can seem more plausible than saying that N acts in Aing just in case N is Aing plus p, since there are some things a subject can't help but act in doing (reading silently, or buying a house), so that there seems to be no room for a separately describable condition fit to make these into actions. By focusing on N's body, we may hold out hope for such an analysis, since there's quite certainly nothing a body can do which on its own *guarantees* that someone is acting in doing anything. (See footnote 13 for hints on why this motivation is problematic.) Another possible motivation lies in the vague thought that an action must *begin* somewhere, and that that somewhere must be in the agent's body, which I don't think is a good thought. (We'll get back to issues about basic action in chapter 4.2.) Still another apparent reason for preferring bodies is a commitment to a form of "causalism" on which causation can only have effects that are bodily movements, and not, for example, instances of waiting. (But it is not clear why a form of causalism which requires that should itself command adoption.)

[...] and the problem [...] is to solve in some philosophically respectable way for x” (Danto, 1981, p. 5). Searle has offered the beginnings of an account of action on which a “successfully performed intentional action characteristically consists [in part in a] bodily movement or state of the agent” (Searle, 1980, p. 47), and does not say much to discourage the impression that he thinks every intentional action consists in a body’s movements. Finally, Michael Smith (2012, p. 387) is a proponent of what he thinks of as the popular thesis that for N to be acting in Aing is for N’s body to be doing something, and for further conditions to apply, as he endorses what he unquestioningly calls “[t]he standard story’s answer [to the question of what acting is]”, which “is that the difference lies in the causal etiology of what happens when a body moves” (Smith, 2012, p. 387). And though I take Smith to represent a larger movement, I will focus on his writings on the topic, as he is an especially clear and steadfast proponent of the present kind of account.

Interestingly, Smith does not only say that for N to be acting in Aing is a matter of N’s body doing something. He also restricts the bodily happenings fit to constitute N’s actions to *movements*. This restriction is, I think, a natural response to a difficulty which the present kind of theory faces. But it brings in difficulties of its own. After a theory like Smith’s is revised so as to avoid both kinds of difficulty, I think it will be fairly apparent what is really wrong with his kind of theory.

To see the temptation of introducing a restriction to movement, start by noting that if we tried to make up a theory saying “for N to be acting in Aing is a matter of N’s body *Aing* and ...”, then we would have a theory which said that for N to be acting in doing N’s taxes is, among other things, for N’s body to be doing N’s taxes. The problem with saying this is not that it is linguistically alien and unpalatable to common sense. The problem is that as long as we

regard N's body as a subject in its own right, separate from N, this claim does not seem capable of meaning something that might be true.¹² (Obviously my objection does not go through if the opposing side identifies N's body as just N — the subject acting in doing the taxes, in this case. But that would mean reverting to the different kind of theory which says that for N to be acting in Aing is a matter of N Aing.)

If we don't want to end up with such disappointments again, we must note that there are only so many things it makes sense to say a body is doing. Bodies, considered as separate from acting subjects, can move, or be still, or sweat, or fall, or burn, and a number of other things, but not do taxes or try to take revenge. Taken on its own, this observation does not require a theorist like Smith to say that a body needs to *move* in action, but such a restriction can seem natural, since (a) perhaps it is thought to provide some informative contribution to an analysis of what acting is, and (b) movement seems to include very much of what a body can do, and because (c) bodies do seem to move quite often when subjects are acting in doing things.¹³

There is a straightforward objection to Smith's restriction to bodily movement, but it seems possible to adjust Smith's kind of theory to accommodate the restriction. The objection is that someone can act in playing dead, which doesn't seem to involve movement, so that an account which tries to understand action through the notion of a bodily movement must fail.¹⁴ Smith's

¹²It does mean something that might be true to say that a body is falling or sliding — the point is not intended to apply for all possible things a subject can be acting in doing.

¹³This kind of theory ignores the complaint that in trying to circumscribe agency, "it will be a mistake to look for *the* fundamental description of what occurs — such as movements of muscles or molecules — and then think of intention as something, perhaps very complicated, which qualifies this" (Anscombe, 1963, p. 29). Though I am sympathetic to this conclusion, it is very difficult to make out Anscombe's argument for it. Some of what I say below is, however, provoked by some of what Anscombe says in that argument.

¹⁴Hornsby (2004, p. 5) makes essentially this criticism, but in a way that saddles Smithesque theories with doctrines about "events" and "actions" which it seems to me these theories don't need to subscribe to, or always do subscribe to. I might, like Hornsby, also have appealed to actions in the category of "omissions": If, for example, I decide to not get out of bed today, not getting out of bed is my action, and nothing is illuminated about it by calling it a kind

own response to this objection is that the notion of bodily movement he has in mind is very broad, and includes things we would ordinarily say don't involve movement. I find this manoeuvre unhelpful. It defends Smith's thesis simply by making it unclear what he means by "movement".¹⁵ But it does not seem necessary to engage in prolonged discussion over what Smith means, or over what "movement" can be allowed to mean, since someone proposing Smith's kind of theory seems free to move beyond that notion. Bodies can't do taxes, but can do more than move. For all I can see, Smith might just as well have claimed: "For N to be acting in Aing is a matter of N's body moving, or being still". This, further, seems equivalent to saying "For N to be acting in Aing is a matter of N's body doing something", if move or be still includes all a body can do. Perhaps Smith's proposed stretched notion of movement is meant to coincide with these broader notions. In any case, this move obviously overcomes the simple objection from still action. But it leaves us with a theory like this:

Smith-inspired: For N to be acting in Aing is for N's body to be Xing, and for a further condition C to apply (where C is a condition concerning a causal relationship between N's psychology and the Xing that N's body is doing).

of "movement". I do not choose this example because I do not want to engage with what I think of as the mistaken view that acting involves movement though omission is a kind of "non-action", which still can be done intentionally (as briefly suggested by Sartorio, 2009, p. 513). Suffice it to say that by "action" I just mean this general category of "things that are done intentionally". Once that category is properly elucidated, there will, I think, be no further need to give an "account" of the subclass of things that are done intentionally and also involve movement, just as there is no special need to give an account of the subclass that involves melting things with the body heat of one's hands.

¹⁵Smith says such complaints rest on an "uncharitable interpretation of what the standard story has in mind when it talks of bodily movements", and suggests the interpretation that "any orientation of the body counts as a bodily movement, even the orientations involved in leaning motionless against a wall, or lying still on a bed, or relaxing on a couch" (Smith, 2012, p. 389). Davidson (2001a, p. 49) has similarly said that we need a "generous" conception of a bodily movement to make good the proposal that actions are bodily movements. But if in a spirit of generous charity, we try to stretch the idea of an "orientation of the body", or of a "bodily movement" to include these passive actions, I do not think we will be left with anything but the notion that the bodies in question are doing *something or other*. Luckily for Smith and Davidson, this claim seems available to them for all they've said.

The resulting story is familiar, though we have left behind the usual restriction to movement: For N to be acting in typing is for N's psychology to cause N's body to do certain (typing) movements, whereas for N to be acting in resting is presumably for N's psychology to make N's body stay put (in the way characteristic of resting).¹⁶ The obvious and familiar problem for the view concerns which causal-psychological relationship an agent is supposed to have to what their body is doing just in case they are performing an action. In answering this question, proponents of Smith's kind of theory must reckon with counterexamples involving deviant causal chains. In the next chapters I will return to questions about how the relation between an agent's psychology and what happens needs to be construed to exclude such counterexamples.

My present objection is different from concerns about deviant causal chains, and similar to my objection in the last section. For it again presses on the claim that if N is acting in Aing, N must be Aing. How can facts of the rough shape that N's body is caused by various psychological states on part of N to do something, entail such facts as that N is doing something? In general it seems that whatever claims we make about what one thing is doing have no bearing on what any other things are doing. If a tornado is passing through, nearby dominoes may or may not be falling over. That a tornado is passing through in such a way that one domino is falling over hardly entails of any other domino that it is falling over, or of any other thing that it is doing anything. So how are we suppose to extract from the fact that a desire makes a body do

¹⁶Since not just any bodily activity can contribute to just any action, such theories must seemingly also introduce some restriction on which movements on part of a body can constitute which actions. To provide such a restriction, a theorist of the present type might try to adapt a suggestion from Davidson (2001a, p. 51) — who is not a theorist of this type — and say that N's body's Xing (moving an arm, say) can constitute N's acting in Aing either by being identical to it (if someone is acting in moving their arm), or by causing it (for example, if someone is acting in communicating some message by moving their arm). It is not clear whether this suggestion can be adapted with much success. (For example, how can we identify a body's activities with a distinct subject's actions?) But here I am ignoring this elaboration and these resultant questions.

whatever movements seem typical of playing the piano that an agent, viewed as something distinct from the body, is playing the piano, or doing anything much at all?¹⁷ (Even unintentionally — since maybe it's a case of desire-triggered sleepwalking.) If, as seems generally true, no fact about what one thing is doing entails any fact about what another thing is doing, an account with the shape of Smith's must fail to provide for the idea that an agent does the things which this account calls the agent's actions.

1.3.3 Agents Doing Actions

A theory of what it is for N to act in Aing can avoid introducing the idea that N is Aing either by wholesale refusal to introduce any subject as the doer of anything that happens, as our first kind of theory did, or through refusal to introduce N as the doer of the things that are N's actions, as the second kind of theory did.

My objections to both kinds of theory hinged on the idea that, if N is acting in Aing, N must be Aing, and on the seemingly necessary failures of these theories in providing for this thought. It seems that, to provide for that thought, we must allow ourselves both to predicate doings of subjects, and, moreover, to predicate them of the N who acts. Once my objection is in view, it can seem obvious: How *could* one provide for the idea that N is Aing in a theoretical mindset where one refuses to predicate Aing of N? What is less obvious is that there are two quite different sources such a denial might have.

Table 1.1 outlines four kinds of account of what it is for N to act in Aing, divided according to whether they do or do not allow mention of N, and to whether they do or do not allow predication of doings to mentioned subjects:

¹⁷With properties, it is normally accepted that they are exclusive of their bearers, so that part of what it is for a thing to be red is for it to be red regardless of which colours or other properties other things have. I am suggesting that we should take the same kind of attitude towards what things are doing.

	Allowing mention of N ...	Not allowing mention of N ...
... allowing mention of what mentioned subjects are doing	N is Aing	N's body is Xing (where Xing \neq Aing)
... not allowing mention of what mentioned subjects are doing	N is F and A is happening	N's body is F and A is happening

Table 1.1: Four ways of restricting answers to the question: “What is it for N to act in Aing?”, with four corresponding types of answer.

The lower-right kind of account might be the one which Nagel targeted to start with. In that case, Nagel’s worries about disappearing agents really had two sources: That the subjects of action tend not to be introduced in accounts of what it is for them to act, and that, in general, the active contribution of substances — the fact that things do things — often fails to be introduced in such accounts. (I am not sure that I can see that both attitudes can be explained by a single commitment to a “third-personal” methodology.)¹⁸

Now my entire critical engagement with the upper-right, the lower-left, and, by extension, the lower-right kind of account, hinged on the idea that for N to be acting in Aing requires N to be Aing, in the sense I have assigned to that latter expression. It might now be questioned whether this is true, since if not I have no case. But I’ll argue that if someone does not affirm that agents do their actions, in the sense that their actions are predicable of them in the sense I started out discovering, their theory must take on a particular relationalist

¹⁸Bishop argues that the “concept of action[...] seems essentially to involve agent-causation” whereas “agent-causation has no place[... in a] ‘scientific image’ of the world [as] a vast array of events and states of affairs related according to laws of nature” (Bishop, 2012, p. 55). If “agent-causation” means that agents do their actions, this brings out, I think, that it is not so much the third-personal perspective, but a certain contemporary view of nature which tries to understand it without the idea that things do things, which generates the second kind of action-theoretic commitment. The first seems to be generated in a different way, which Nagel and some of his followers sometimes gesture at: The things we want to identify as the sources of action (which are not mere bodies) disappear when we view the world third-personally. Whatever the reason is for which these subjects are thought to disappear, it seemingly applies *whether or not* we adopt the just-ostended “contemporary” view of nature.

shape. And it will be the subject of the next chapter to undermine prominent instances of that form of theory.

1.4 Two Approaches

I argued that, if someone gives an account of what it is for a subject to act in doing something, which does not introduce the idea that the subject is doing the thing in the sense which I started out presenting — the sense of progressing towards success or failure in doing it — then this theorist will find themselves unable to secure that idea by talking, for example, about what an agent's body is doing, or about happenings which are not thought to be doings on the part of a specific subject.

But what if a theorist would opt to look for an account of what it is for *N* to act in *Aing*, without committing to affirming that this requires *N* to be *Aing* in the sense I have sketched? If a theorist does not take on such a commitment, they seemingly cannot be criticised for failing to do justice to it. But then such a theorist cannot take the question “what is it for *N* to act in *Aing*?” as presupposing that *N* is *Aing*, as I have taken this question to do. They must then construe this question as bearing a somewhat different sense.

How might such an action theorist get their inquiry off the ground? What question could they seek to answer, if not one which presupposes that a subject's actions are predicable of this subject, in the way that other doings are? Though I have no indisputable argument for the following conjecture, it seems that the main alternative approach — perhaps the only stable one — is as follows: First, observe that there are things that happen. Second, conceptualise these as events, where perhaps this means that they are individuals, as Davidson thought. Now pose the action-theoretic question in this way: What is the *relation* between an

agent and an event which makes it into their action?¹⁹

Inspired by the early Davidson, a number of philosophers have pursued such questions, attempting to specify the nature of action by presenting a special kind of relation thought to be exemplified only and wherever some event is an action. Hence such a theorist attempts to understand what it is to act, not by predicating actions of subjects, but by introducing some or other *two-place predicable*, relating a subject and their action, now considered as an individual.

Or, as sometimes happens, such a theorist might abandon or disregard the Davidsonian assumption about events being individuals, thereby introducing something like a relation (if indeed this makes sense) between a psychological *fact about a subject* and a *fact about what they are doing*. Since it is the invoked relation that is supposed to do the work in explaining the nature of action, such a theorist can seemingly remain non-committal about just how to understand the fact about doing figuring at the far end of the invoked “relation”.²⁰

¹⁹The original motivations for thinking of actions as individuals related to subjects are not overtly action-theoretic, and are no longer in much focus. One important motivation for viewing actions as events, where these are considered as particulars, stems from Davidson’s concern with providing an account of the structure of sentences like “Sebastian strolled through the streets of Bologna at 2” on which it can be shown they entail ones like “Sebastian strolled through the streets of Bologna”. Davidson says the real structure of the first sentence is that “There is an event x such that Sebastian strolled x , x took place in the streets of Bologna, and x was going on at 2” (Davidson, 2001a, pp. 166-167). Based on further comments (Davidson, 2001a, p. 175), it seems that we are here supposed to view “Sebastian strolled x ” as introducing Sebastian, a two-place predicate (of strolling), and an event (a stroll). Henceforth I will not be concerned with stating or criticising the kind of conception of events as individuals which would be needed to make sense of this semantic suggestion. But just to say something about why Davidson’s line might lead to a blind alley, I want to point to the oddness of thinking that, whenever we would ordinarily say and think that a subject is doing something, this is really a matter of the subject and an event (for which we can find no clear application for the question “which one?” — an embarrassment that Davidson (2001a, pp. 309-310) acknowledges) being related through some two-place predicable. This suggestion is somewhat analogous to a suggestion which the reader might more readily recognise as degenerating: That “this thing is red” specifies of the thing mentioned that it is related in a special way to a special individual, namely redness. Just as we want to ask, of Davidson’s account, what an event (considered as an individual) is supposed to be, and how it is supposed to be related to a subject just in case it is their action, we want to ask, of the present account, what a colour (considered as an individual) is supposed to be, and how it is supposed to be related to a subject just in case it is red.

²⁰The issue of just what can figure as a relatum in the kind of relation that such an action theorist needs is very large. For an extensive discussion, see Stein (2014). I suspect, however, that the discussion in the next chapter is going to be neutral on the question of whether relations, and particularly causal relations, should be thought to hold between individuals or

Either kind of relationalism is to be contrasted with a form of account which rather views a subject's actions as directly predicable of this subject, in the same way as any doing is predicable of the subject doing it, and which must hence find a different way of explaining the difference between acting and merely doing. The next chapter is devoted, first, to criticising the predominant form of the above form of relationalism, and, second, to working out an account of this different shape, immune to the illnesses that affect such forms of relationalism.

something like facts.

Chapter 2

Doing What's Wanted

2.1 Introduction

This chapter has a negative part and a positive part. The negative part goes as follows: A family of prominent accounts endeavour to explain the nature of acting in terms of a specific kind of causal relation between wanting and acting. Each account must say more than this about the nature of acting, since deviant causal chain cases show that the supposed kind of causal relation can exist where no one acts. Each account tries to introduce an extra condition to make up for this difference. All fail, since each faces a version of the following dilemma: Either the extra condition is simply understood as that which makes up the difference between deviance and action, in which case invoking it is not informative, or the introduction of the extra condition makes the proposed account come out false.

The positive part of the chapter is devoted to reformulating the problem of deviant causal chains as an instance of a more general problem of distinguishing exercises of dispositions from happenings which mimic such exercises. The

chapter warns against a false start which attempts to cut out deviant cases by introducing a generic conception of exercising dispositions, while otherwise clinging to something like *Standard Causalism*. It recommends, instead, that each kind of disposition that might be exercised needs to be understood through individuating the specific kind of condition in which it is exercised. In line with this recommendation, the chapter outlines an account of acting as exercising a kind of disposition to act in conditions where the acting subject has a belief which represents this subject as acting.

2.2 Actions as Causal Relata

I take for granted that Velleman is right to say that the following outlines a popular, causalist theory of acting:

There is something that the agent wants, and there is an action that he believes conducive to its attainment. His desire for the end, and his belief in the action as a means, [...] causes the corresponding movements of the agent's body. (Velleman, 1992, p. 461)

This outline of a theory says that acting is what happens when (a) a subject “wants something” such that the subject believes something “conducive to its attainment”, and (b) these states (together) “cause” something to happen, which (c) “corresponds” to these psychological items.

What Velleman says is a mere outline because it does not convey anything very determinate about (a), (b), or (c). A more determinate understanding of (a), (b), and (c) is, however, implicit in the popular theory Velleman outlines, and which I will call “Standard Causalism”. For theorists who endorse Velleman's formulation do tend to have more specific conceptions of at least (a) and (b), which, whether or not theorists explicitly think so, force a particular con-

ception of (c) too. In the following three sections, I make explicit these ways of understanding (a), (b) and (c) that are implicit in *Standard Causalism*.

2.2.1 Psychology

On the nature of wanting something to happen such that one believes something conducive to its attainment, we can comfortably allow that standard causalists are divided along many axes,¹ as long as we see that they must all endorse the following: Whatever it is to want something and believe something conducive to its attainment, it is to be in some state which, first of all, can potentially cause something to happen, and which, second of all, can correspond to what happens. This much is needed just to consistently maintain the rest of the standard causalist theory. For simplicity, and to maintain verbal compatibility with others who have introduced the same simplification, I will henceforth often refer to this state of wanting to do something such that one believes something conducive to its attainment simply as a state of wanting to ..., making the reference to an instrumental belief about what's wanted tacit. For further simplicity, I will not engage in any of the possible and actual disagreements about how a standard causalist should conceive of these psychological states. So I will not wonder, for example, whether they are states of the brain, or whether, if not, they must be conceived dualistically. My upcoming argument will just concern the relation between such states and what happens, which *Standard Causalism* thinks makes these events actions.

¹One axis of division is the question of whether the desires required for action constitute a belief-distinct kind of attitude (Smith, 1987, pp. 55-57) or are just a special kind of belief (Setiya, 2007, pp. 49-51). In the main text I assume that I can stay neutral on this question. If these states are not distinct, it would be unavoidable to refer to believing while desiring as a single state. But even if they are distinct, it still seems possible to think of them in conjunction as making up a single state, analogous to someone's (single) state of believing that p while believing also that q.

2.2.2 Causation

Then what is it, according to *Standard Causalism*, for a desire to cause what happens? Davidson famously appealed to “the ordinary notion of cause” in stating his claim that desires cause actions (Davidson, 2001a, p. 9).² Disregarding “ordinary”, calling it *the* notion suggests that the appeal is to some one general notion of something making something else happen, so that whatever it is for a desire to cause an event, in the sense that the theory requires for acting, it is the same sort of thing as is exemplified when the weakness of a branch makes a sloth fall, or when the attack of a bowling ball breaks a chair, or when a mistake concerning ginger determines which words appear in a restaurant’s review.³ General as it is, this notion of causation must seemingly be applicable also when someone’s desire makes something happen which is *not* an action, as when someone missteps, shakes, or feels sad because they have a desire that is shameful. Though this supposed idea of causing is still obscure, we know something about which extension it must have: It applies not just where desires cause actions, since it can apply also where non-desires cause non-actions, and where desires cause non-actions.⁴

²In the main text of this section I do not present the motivations for this view. One commonly stated motivation is that only the present kind of causal theory will allow us to reduce actions to a scientific worldview which requires that everything be explicable in terms of such causes. This motivation seems quite weak to me as it stands, since it is quite hard to bring out why a scientific worldview that requires the present kind of causalism should itself be required. In the next footnote, I gesture at a different, and I think more pressing motivation, due to Setiya. In the following section, I present a close cousin of Davidson’s original argument for the claim.

³In phrasing things in the above sort of way, “a desire makes ... happen”, etc., I perhaps suggest an answer to Davidson’s (2001a, pp. 151-162) and Stein’s (2014) questions about whether causers and causees are individuals (desires, events, etc.) or something like facts (that N wants to ..., that ... happens). Much of my language suggests the former over the latter, but, as was hinted in note 20 on page 32, if it is possible to work out an alternative conception according to which causation is a “relation”, of sorts, between facts, not individuals, it seems that my upcoming discussion, including the criticisms, will apply equally to that conception.

⁴Setiya explicitly endorses a claim which, I think, comes down to a version of the present one, when he says that “what it is to act for a reason is a matter of psychological causation, in a sense of ‘cause’ that also applies to causal deviance” (Setiya, 2011b, p. 131).

2.2.3 Correspondence

Finally, what, on *Standard Causalism*, is it for a desire to correspond to what happens? The basic and vague idea must be that acting requires that the state of wanting to ... represents the kind of thing that happens, since seemingly someone can't act in watering flowers if they don't have psychological states that represent what they are doing as a case of that.⁵ But how, and how strongly, should we conceive of this notion of representing "the kind of thing that happens"? There seem to be two interesting options, one weaker and one stronger:

Weak Correspondence Condition For someone to act in *Aing*, it is necessary that this someone has a desire that represents N as *Aing*

Strong Correspondence Condition For someone to act in *Aing*, it is necessary that this someone has a desire that represents N as *acting in Aing*⁶

Importantly, the weak correspondence condition is satisfied even in some cases where a subject doesn't act in *Aing*. For Someone can have a desire which represents them simply as watering flowers, and be doing that, and still not act in watering flowers (— perhaps there's no water left in the heavy vase they are using but they're unwittingly sweating an implausible amount onto the flowers in lifting it).

The strong correspondence condition does not leave room for this possibility. For it says that subjects who act represent themselves as *acting*, and not just as being engaged in activities in the way that non-agents can be engaged in activities. Because of this, people who do not act cannot satisfy the condition (with respect to the specific non-action in question): If one doesn't act in watering

⁵This condition on action is of course crucial to Anscombe (1963) and to such followers of Anscombe as Setiya (2007).

⁶The second condition is stronger since representing oneself as acting in *Aing* entails representing oneself as *Aing*.

flowers (as in the sweating example), psychological states that represent one as acting in watering flowers can't correspond to what happens.

Which of these conditions is intended by *Standard Causalism*? This question is generally ignored by those who endorse the theory. This might make us think that theorists are free to endorse either. But in fact I believe it can be shown that *Standard Causalism* lacks motivation on the stronger condition, so that reasonable standard causalists must have in mind something like the weaker condition. To see this, consider what, to a standard causalist, is supposed to motivate introducing a *causal* condition alongside the correspondence condition. An argument congenial to Davidson's classic complaint against pure "correspondence theories" of action suggests itself:⁷

A person can *A*, and have a desire that corresponds to what they are doing, and still not act in *Aing*. Because of this, there must be some condition, apart from *Aing* and having such a desire, that is necessary for a acting in *Aing*. Our best bet for specifying the missing necessary condition is as follows: It is necessary for acting in *Aing* that a desire which corresponds to what they are doing also *causes* the agent to *A*.

Ignore the tempting question over whether this argument is valid.⁸ Note instead that, if the premise with which it starts is false, it can't support its conclusion. Then note, as I have just argued, that on our second conception of correspond-

⁷The argument I have in mind is the one that starts with the premise that "a person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it", and his subsequent attempt to distinguish acting *for* the reason had by appeal to the "idea that the agent performed the action *because* he had the reason" (Davidson, 2001a, p. 9). It is somewhat reinforced by the subsequent suggestion that this introduction of a causal condition "alone promises to give an account of the 'mysterious connection' between reasons and actions" (Davidson, 2001a, p. 11).

⁸For brief presentations and dismissals of some non-causalist accounts of the "mysterious connection", see Davis (2005, pp. 79-80). For what seems to be a quietist "account", see Dancy (2000, p. 163). Setiya (2011b, p. 144) briefly questions Davidson's inference but suggests a different route to its conclusion.

ence, the premise is false. We must infer that, on the second conception of correspondence, the argument doesn't support its conclusion.

Hence in so far as a theorist wants to rest their standard causalism on an argument like the above, which crucially mentions the inadequacy of the correspondence condition, they cannot also appeal, in their standard causalism, to the stronger conception of correspondence, which crucially makes such correspondence sufficient for acting. Such a causalist must hence intend either the first conception of correspondence, or (at least) some intermediate conception of correspondence which still allows that desires can correspond (in the intended sense) to what happens when what happens isn't a case of acting.

Since a standard causalist thinks that causation by desire is insufficient for acting (since they appeal to a general notion of causation), and since a standard causalist introduces causation by desire on the basis that representation by desire is insufficient for acting, such a theorist must think, at least, the following:

Standard Causalism If N acts in *A*ing, then N has a desire that represents N as *A*ing (in a sense such that N can *A* and represent N as *A*ing without N's acting in *A*ing) and this desire causes N to *A* (in a sense such that this desire could cause N to *A* without N acting in *A*ing).

And this formulation adds some determinacy to Velleman's characterisation of *Standard Causalism* as a theory which defines action in terms of correspondence to and causation by desire.

But of course a simple conditional does not give us an account, but just implications of acting. A first pass at providing an account would be to simply replace the above conditional with a biconditional.

Naive Standard Causalism N acts in *A*ing *just in case* N has a desire that represents N as *A*ing (in a sense such that N can *A* and represent N as *A*ing without N's acting in *A*ing) and this desire causes N to *A* (in a sense

such that this desire could cause N to A without N acting in Aing).

It straightforwardly emerges that on *Naive Standard Causalism*, acting is conforming to two conditions, each of which, on its own, is insufficient for acting.

Below I will argue that *Naive Standard Causalism* is undermined by cases involving “deviant causal chains” (hence I call it “naive”). Going further, I will argue that attempts to preserve the heart of the account by introducing qualifications to the right of the above biconditional fail, since these qualifications will either point uninformatively right back to the idea of acting for their sense, or simply make the resulting account false.

2.2.4 Hopeless Deviance Problem

The standard counterexample to *Naive Standard Causalism* employs the idea of a “deviant causal chain”. Here is as good an example as any of a deviant causal chain:

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, and yet it might be the case that he never chose to loosen his hold, nor did he do it intentionally. (Davidson, 2001a, p. 79)

Davidson’s climber falsifies *Naive Standard Causalism* since this climber has a desire that represents letting go, and causes letting go, in just the way *Standard Causalism* understands these conditions, but still this climber does not act in letting go.⁹

⁹I ignore the possible but unpopular suggestion that Davidson’s nervous mountain climber does act in letting go. Claiming that seems analogous to claiming that people in Gettier cases know what most people would say they truly but accidentally believe. Of course the unpopularity of these suggestions doesn’t undermine them, so that one might still wonder how

The only conceivable way of improving on *Naïve Standard Causalism* while still building on the foundation of *Standard Causalism* is by introducing additional conditions with the aim of excluding cases like the mountain climber while still including all genuine cases of acting. There has not been a shortage of attempts, and below I present three prominent ones. I argue that a single and fairly mechanical procedure undermines each, which just involves asking: “What is it for the extra qualification to apply?” and then noting how each account seems forced to answer this question either by saying something that is uninformatively circular, or something that is not circular but clearly fails to separate acting from deviantly doing.

Before proceeding to discuss these attempts at completing *Standard Causalism*, it is perhaps worth engaging with a kind of quietist view which says that there is a mode of causation that is definitive of the desire-caused events which are actions, without thereby committing to specifying what constitutes that mode of causation. Setiya has at one point proposed that such a retreat is open to a standard causalist. For before going on to attempt a specification of the right kind of causation, he has said this:

[N]othing I say below turns on the claim that we can explain what “deviance” is in an illuminating way[...]. My point is *metaphysical*: the intentionality of action consists in psychological motivation [...], and motivation consists in a certain kind of efficient causation. It does not matter whether we can explain this kind of causation to someone who lacks the concept of intentional action. (Setiya, 2003, p. 348)

they could be undermined. Though I do not pursue it, I imagine that there is an argument to be made against both of the above suggestions which involves questioning whether the requisite ideas of acting or of knowing can have the connections with related ideas (including a connection with the idea of acting for a reason in the case of acting, and the connection with the idea of being justified in believing in the case of knowledge) without which they seemingly would be quite different and alien notions of “acting” or “knowing”.

For the purpose of showing further examples of this kind of retreat, we may note, as Setiya hints, that Davidson can seem to go along with some such thought, on the one hand committing to the “incomplete and unsatisfactory” account that “an action is performed with a certain intention if it is caused *in the right way* by attitudes and beliefs that rationalize it” (2001a, p. 87, emphasis added), while on the other hand expressing “despair of spelling out [...] the way in which attitudes must cause actions” in such cases (2001a, p. 79). (But Davidson’s rueful wording signals that, even if he contemplated such a retreat, he did not consider it entirely satisfactory.¹⁰) Goldman might be taken as a clearer proponent of Setiya’s kind of retreat, since he endorses an account in the standard style while saying that he doesn’t “think it is fair to demand of a philosophical analysis that it provide” action’s characteristic mode of causation (1970, p. 62). (But as Goldman suggests that the relevant analysis might one day be provided by scientists, he too apparently registers some need to provide such an analysis.)

Perhaps Goldman, Davidson and Setiya have never whole-heartedly proposed such a retreat. Despite what’s said in the above quote, Setiya has always endeavoured to state what makes causation non-deviant, and anyway did not keep the above passage when adapting the associated paper for a book (Setiya, 2007). But it is worth bringing out what makes such a retreat unsatisfying, so as to remove the appearance that it is optional for a standard causalist to try to spell out what non-deviant causation amounts to.

As Setiya presents the proposal, the idea in it seems to be to say that a specific kind of causation is involved only in those cases where subjects act,

¹⁰Davidson’s later stance seems more clearly pessimistic about *Standard Causalism*, saying that “the concepts of event, cause, and intention are inadequate to account for intentional action” and apparently being unmoved by other philosophers’ attempts fill in the gap using additional notions (2004, p. 106). In the light of this later pessimism, what’s said in the main text may be viewed as a transient intermediate development in Davidson’s thinking: A wish to endorse something like *Standard Causalism* coupled with inability to see how such a theory could be brought home. Or it may be viewed simply as a somewhat roundabout way of saying that a theory along standard causalism’s lines can provide no account of acting, though it can provide necessary conditions for acting.

without saying what that kind of causation is. The proposal seems to rest on some prior and general way of dividing causation into kinds. If we take for granted that there is some clear way of doing this, the proposal still does not tell us which kind is pertinent to action. If, instead, we suppose that we're in the dark about how to distinguish kinds of causation, the proposal becomes hard to distinguish from one according to which there is *some condition* or other such that, when it obtains, causation by desire amounts to action. In either case, the proposal seems to come to giving an account of acting, which claims some extra condition is required for acting, but to then say that no specification of the extra condition is needed, and that the account is still in order.

But how could an account be in order if it merely appeals to the existence of some extra condition, without endeavouring to state the condition? Developing the hint in the quote from Setiya, it might now be suggested that there are two kinds of account, one metaphysical, and another which we might dub “conceptual”. It might then be suggested that a conceptual account must be capable of explaining its target concept to someone who lacks it, whereas a metaphysical account need not be capable of this. Finally it might be suggested that if the present account of acting is merely metaphysical, then because it is not subject to the demand of being capable of providing someone with its target concept, it escapes such complaints as I tried to press — complaints of being an inadequate account.

Now there are two ways of developing Setiya's hint further. Assume that a metaphysical account need not be capable of explaining its target concept to someone who lacks it. Must such an account be capable of explaining its target concept at all? If it does, then we should consider whether the account which Setiya seems to offer does provide explanation. It does not seem to do so. For even when granting that a metaphysical account need not be capable of

explaining its target concept to someone who lacks it, the presently proposed account does not even seem capable of explaining its target concept to someone who has it: I think I know what acting is, but I do not think what Setiya says allows me to understand the idea there's a special kind of causation which never exists in deviant cases, and always exists in action — indeed I doubt that there's such a special kind of causation, if causation is construed along relational lines.

Now assume, instead, that a metaphysical account need not be capable of explaining its target concept. It might be that when Setiya says that it “does not matter whether we can explain this kind of causation to someone who lacks the concept of intentional action” (Setiya, 2003, p. 348), he takes it as given that, if we can't explain it to someone who lacks the concept, then we can't explain it at all, since the only respectable kind of explanation of a concept is of the sort that could equip the explainee with it. This assumption is very large, since it seems to amount to the thought that an explanation of a concept must take a reductive route — must show a way of constructing the concept out of independently available materials.¹¹ But the assumption is also very troublesome when considering the present suggestion for a merely metaphysical account. For if it is assumed that a philosophical account, when it is merely metaphysical, need not provide any insight, elucidation or explanation of its target concept, then what reasonable standard could there be by which to measure it, and what could be the point of providing a merely metaphysical account? The thought that explanation need not be involved in metaphysical accounts makes it very hard to understand the idea that they make up a species of account.

All I want to suggest by the previous is that there does not seem to be a stable way of maintaining an account without trying to state all the conditions which this account introduces. Hence a standard causalist must take a more ambitious line than the above and essentially quietist one: They must attempt

¹¹I will return to this question, in a way, in chapter 5.3.

to say what makes an action nondeviant, and may not just say that *something* does. Next I discuss some prominent attempts to do this.

Begin with Setiya’s own suggestion, that in action, a desire “not only causes but continues to guide behaviour towards its object”, and that “[i]t is this condition that fails in Davidson’s example” (Setiya, 2007, p. 32). The suggestion obviously raises the question: What is it for a desire to continuously guide behaviour in the intended sense? An uninformative answer would be, for example, that “for a desire to continuously guide behaviour is for the desire to take part in the production of an event in such a way that that amounts to *acting*”. There are variations of this answer, which still are uninformative, including more convoluted answers which specify continuous causation as the kind of causation that is “characteristic” of actions and not of cases like the mountain climber’s. They are uninformative because they do nothing to identify the required kind of continuous guidance, except as a feature which is a mark of actions and not of deviant cases.

Here is a different kind of answer, which is informative: A desire continuously guides what happens just in case no break or pause is involved in the unfolding of the event. This is no good, since we can easily imagine a version of Davidson’s mountain climber who deviantly lets go but does so quickly and smoothly, so that there is no such break or pause, and since we can imagine versions where similar pauses (hesitations, momentary stumblings, slips and mistakes) occur though the climber does act.¹²

¹²If someone should respond that there is a specific kind of break that is always present in deviant cases, so that it is a mark of acting that there is no break of that kind, then something elucidating must be said about this kind of break, which does not invite the same objection about unformativeness as was just raised. One problem for such a response is that, when there is a type of doing, A, acting in doing which requires that there is no interruption, pause, breakdown or hiccup of type X, there will very often be another type of thing, B, acting in doing which requires that there *is* a break of type X. To illustrate this, we may suppose that there is a kind of break absence of which is a mark of a climber’s acting in letting go. Maybe it is necessary that the rope doesn’t unexpectedly slip out of their sweaty hand. We can easily see that a saboteur could be acting in ensuring that the rope does slip out of the climber’s hand in just the way considered (by greasing the rope or taunting the climber, perhaps). This

Next there is a proposal of Mele's (1992), which introduces "proximate causation" in place of Setiya's "continuous guidance". Sehon's (1997, pp. 199-202) persistent attempts to clarify this notion have shown that it encounters a similar dilemma between unformativeness and falsity.

For to decide whether a desire proximately causes an event, say, of letting go of a rope, we need some criterion for whether the things intuitively happening "in between" the desire and the letting go, such as nervous shakings, constitute a separate and interfering event or are part of the agent's letting go. What then is it to be a separate and interfering event? We could of course say that interfering events are those that do not contribute to, or are not part of, the agent's *action*. But again, this move would make it uninformative to say that such proximity is required in action.

On the other hand, we may try to decide by some independent criterion whether a desire proximately causes what happens. Mele (1992, pp. 201-202) apparently makes this simple suggestion: For some bit of psychology, A, to proximately cause an event, C, is for A to initiate the "physiological chain" that constitutes C (Mele, 1992, pp. 202). This formulation makes obvious the problem with Mele's proposal, (observed by Sehon, 1997, p. 201), which is that for all we know about what it is for an event to initiate some physiological chain, this condition does not seem to ward off the deviant outcomes we have been concerned with. If a desire causes nervousness which causes, for example, some bit of bodily movement, then there is every prospect for claiming that the nervousness and the movement together constitute a "physiological chain", initiated by the desire, and Mele's observations on the matter do little to tell us

seems to show that, to work, the present response needs to relativise the kind of break that makes for deviance to the thoughts or plans of the acting subject. But that would seem to require a departure from *Standard Causalism*, since it would seemingly amount to introducing a notion of acting according to plan, or of correspondence between thought and action, which itself excludes deviant cases.

why they should not so count.¹³

I now move on to a more sophisticated way of accounting for deviance, which, on closer inspection, falls prey to the same simple form of criticism. Authors like Peacocke (1979), Bishop (1989) and Smith (2012) have said that in deviant cases, there is something accidental about the causal connection between desire and event. And they have tried, to various extents, to explain the contrast between deviance and action as a contrast between what non-accidentally happens and what accidentally happens. Peacocke has gone furthest in trying to express what the requisite notion of non-accidentality comes to. He has emerged with the notion of *differential explanation*. If I understand Peacocke's notion of differential explanation, and the way in which it is supposed to apply to action, his proposal is as follows (and I will ignore that Peacocke prefers stating his theory in terms of intentions rather than desires, since in this context nothing seems to hinge on their possible difference):

Differential Explanation For a given desire to differentially explain a given bit of activity is for there to be a true law of nature which poses a suitably robust explanatory relationship between a kind to which the desire belongs and a kind to which the activity belongs. (Peacocke, 1979, pp. 63-71)¹⁴

Peacocke's discussion of what makes for such an explanatory relationship between kinds is quite involved, and therefore hard to readily assess. The rough idea seems to be that two kinds stand in this robust explanatory relationship just in case there is a true law of nature which contains a function, in mathematical

¹³Sehon goes further than I do in engaging with the possible epicycles that might be tacked on to Mele's suggestion, including prohibiting a list of abnormal circumstances, requiring "normal circumstances", and requiring "continuous causation" (Sehon, 1997, pp. 201-203). The first two ideas seem fairly obviously ineffective in their own right, meeting with standard objections about the infinity of the requisite lists, and the futility of attempting to say what normal circumstances are. The last I have already criticised.

¹⁴See Sehon (1997) for what I think is the clearest discussion of Peacocke's notion of differential explanation. My understanding of Peacocke is greatly indebted to Sehon, but as will be seen below I will endeavour to provide a criticism different from Sehon's, targeting the stage in Peacocke's theory where he introduces the notion of realisation.

style, to the effect that, if there is an occurrence of the first kind, there is, or will be caused to be, one of the second (Peacocke, 1979, p. 66). There are more shades and puzzling features to Peacocke's view than this, but it is, in fact, not necessary to treat the notion of differential explanation at much length, since my upcoming criticism does not concern whether desires differentially explain the events that are actions. For in fact Peacocke does not make this claim, but makes the relationship between differential explanation and causation by desire more complicated. This further bit of complication is my target.¹⁵

For imagine that, on discovering the differential explanation condition, Peacocke were to suppose this:

Causalism plus Differential Explanation N acts in *A*ing *just in case* N has a desire that represents N as *A*ing (in a sense such that N can *A* and represent N as *A*ing without N's acting in *A*ing) and this desire causes N to *A* (in a sense such that this desire could cause N to *A* without N acting in *A*ing) *and* that N has this desire differentially explains that N is *A*ing

The problem with such a proposal is the Davidsonian commonplace that there is no suitably robust explanatory relationship which could make the above thesis true. For example, there does not seem to be any robust explanatory relationship between wanting to turn on a light and a light being turned on, because of the multifarious opportunities for weakness of will, conflicts of interest, confusion, lack of knowledge about available means, not to speak of physical malfunction, sudden death, simple temporary disinterest, or procrastination. The results of this seems disappointing: Even if the condition that N's desire differentially explains what N does could rule out deviance (a claim we have not

¹⁵An advantage of this criticism is that it would seem just as applicable if Peacocke had gone for some other account of the sought-after, non-accidental dependence of action on desire. So the criticism would have applied just as well if Peacocke had attempted to account for such non-accidental connections by saying that in them, the occurrence of an action *counterfactually depends* on the presence of desire, or if he had said that actions but not deviant outcomes are *predictable* by appeal to the presence of a desire.

even considered), it could not figure in a satisfactory account of action, since the resulting account would make the relationship between wanting and acting more mechanical or sure-fire than it plausibly is.¹⁶

The points in the previous paragraph are familiar to Peacocke. They seem to be the reason why he adds the following twist to his theory: It is not such facts as that N wants to A that must differentially explain N's Aing when N is acting in doing it, but rather such facts as that N is in some *physiological state*, which in turn *realises* N's desire.¹⁷ And for a brain state to realise a desire is for the desire and the brain state share causes and effects (Peacocke, 1979, pp. 116-124). Hence in so far as Peacocke's account is meant to solve the problem of deviant causal chains, it seems fair to represent it as follows:

Peacocke's Causalism N acts in Aing *just in case* N has a desire that represents N as Aing (in a sense such that N can A and represent N as Aing without N's acting in Aing) and this desire causes N to A (in a sense such that this desire could cause N to A without N acting in Aing) and there is a brain state which differentially explains N's Aing, *and* this brain state shares causes and effects with N's aforementioned desire

Several theorists have queried the idea of differential explanation that Peacocke introduces, but I think a clearer objection can be made by shifting attention to the idea that a brain state *shares causes and effects with a desire*. For suppose that we have isolated some brain state which — whatever precisely this amounts to — stands in Peacocke's differential explanation relation to some event. We then immediately face an issue which seems to come to an insurmountable problem: How is Peacocke's theory supposed to tell us when a brain state shares

¹⁶Of course someone might endorse a version of Peacocke's claims that does not model the requisite robust dependence of action on desire on a mathematical function. But the problem is that we have no clear grasp of such an alternative model, much less of how it could contribute to solving the problem of deviant causal chains.

¹⁷See Bishop (1987); Sehon (1997); Peacocke (1979, pp. 69-70) for further discussion of realisation.

causes and effects with a desire, so that we can go on to declare that this theory has the result that the event is an action? Peacocke does not seem to give us a clear answer to this question.¹⁸

I believe the reason why Peacocke does not answer it is that he can't. For this question raises the previous kind of dilemma again, although in a more roundabout way. On the one side, we could say things like "the brain state shares causes and effects with a desire when the event it differentially explains is an *action*", or more convolutedly, "the brain state shares causes and effects with a desire when it shares causes and effects with the kind of psychological state such that, when it causes an event (and other conditions, like the differential explanation condition, hold), the event is an action", but these answers do not contribute to an informative account of non-deviance.

The alternative is to come up with an action-independent way of getting hold of desires, and to check whether any desires we thus get hold of share causes and effects with Peacocke's differentially explanatory brain states. Our options are many. We could identify desires, with Lewis (1966; 1972), as theoretical entities that fit folk-psychological platitudes about desiring. We could identify them, with Humeans like Sinhababu (2009), as a kind of felt emotions. Or we could identify them, with functionalists, as those states of the organism which tend to produce certain outcomes on certain stimuli (Levin, 2013).

The list of possibilities obviously goes on (including combinations of the above and any further proposals), but it does not seem to matter which possibility we pursue. For it seems perfectly possible for there to be cases where a brain state differentially explains what happens in Peacocke's sense (whatever precisely differential explanation entails), and where, in addition, this brain

¹⁸At a key moment in laying out the idea of realisation, Peacocke says that "any causes or effects of a human being's having a belief, desire or intention will be causes and effects of his brain's being in a certain state" (Peacocke, 1979, p. 118). This says nothing about what it means for some desire to share causes and effects with some brain state, and thus it does not answer our question, but presupposes an answer to it.

state shares causes and effects with any such independently specified “desires”, but where, still, what happens does not amount to acting.

For example, we could imagine that Davidson’s mountain climber lets go of a rope, and a brain state differentially explains this, and that brain state perfectly realises some emotion (by the standards of Humeans, the felt emotion is exactly like that felt when wanting to act), or that it perfectly fits the folk-psychological profile for desiring to let go (by the standards of the folk, the behaviour is just what they expect of someone “wanting” to let go of a rope), or that it perfectly instantiates a functional state thought to realise such a desire (the agent is in just the sort of state — a functional state of nervousness, as it were — that would tend to produce such an outcome in the present kind of condition), or — why not — that all of the above conditions hold, as we might have done all along, depending on how we imagined Davidson’s mountain climber. The climber still might not act in letting go of the rope.¹⁹

What can be learned from the fact that Setiya’s, Mele’s and Peacocke’s accounts fail in similar ways? There is a temptation to argue from the fact that a single and fairly mechanical kind of objection seems to undermine these accounts, to the conclusion that it is impossible to come up with a successor account which avoids the objection. But though the failures are suggestive, no such general pessimistic conclusion seems to follow from the failures of a few attempts, nor does an inductive argument along the same lines seem convincing.²⁰ In the end I am not sure how to give any sort of conclusive argument

¹⁹Sehon (1997, p. 211) notes, I think, that Peacocke doesn’t do anything to show this sort of case to be impossible.

²⁰Wilson attempts to argue directly from the failures of extant versions of *Standard Causality* to the failure of any possible successors. His argument can seem *inductive*, since he says that “the evidence points to more than infelicity or incompleteness in the various causalist proposals – it points, that is, to a global breakdown in the whole project of reduction” (Wilson, 1989, p. 258). One weakness of such an argument is that any reasonably short discussion is bound to introduce only a few causalist proposals, so that we might say that the inductive base is weak. Another weakness is that even a big inductive base would give meagre reason for favouring the “hypothesis” that there are zero true versions of *Standard Causality* over the hypothesis that there is one, which of course is all that any sensible standard causalist

for such pessimism.

So I propose, instead, to challenge those who think that there must be some such account. Why think that? Given the failures of extant proposals, I do not know of a better argument than one which proceeds, somewhat in Davidson's style, from a lack of good alternatives. But this means that the existence of a good alternative would leave us little reason to believe in a second coming of *Standard Causalism*. This makes it seem dialectically unavoidable to double check whether a satisfactory alternative couldn't be developed. So let us continue building on the alternative suggestion from the first chapter, which does not proceed by invoking relations, but tries, instead, to explain the difference between those doings which are actions and those which are not.

2.3 Acting as Exercising a Disposition

As Frankfurt (1978, p. 162) has pointed out, the question of what it is for a subject to act, as opposed to merely doing something which is similar to acting (as happens in the mountain climber example), seems analogous to the question of what it is for a spider to move its legs, as opposed to doing something which is not but mimics such leg movement (as happens if someone manipulates its legs with strings). Frankfurt does not do very much to answer either question, but suggests that, since both concern contrasts between "instances in which purposive behavior is attributable to a creature as agent and instances in which this is not the case" (Frankfurt, 1978, p. 164), both questions ought to be answered by appeal to some general way of elucidating such contrasts.

Though Frankfurt is not explicit on this point, there seems to be a multitude of further examples of the same kind of question, concerning the same kind of contrast. Setiya (2011b, p. 137) suggests several questions which seem to be

will claim.

on a par with Frankfurt's, including the question of what it is for a flower to bloom, as opposed to doing something which isn't but mimics blooming (as happens if something forces its petals to open).²¹ Despite Frankfurt's talk of "creatures" and "purposiveness", such questions do not even seem limited to the activities of life forms. For as Alvarez and Hyman (1998, pp. 243-245) emphasise, some seemingly analogous contrasts seem to apply to the activities of the inanimate, inviting such questions as what it is for iron to rust, as opposed to doing something which isn't but mimics rusting (as I suppose would happen if, through some odd process, bits of iron get gradually replaced by bits of rust).

Frankfurt's brief pronouncements suggest that some one general strategy might be developed to explain how such things as acting, moving, blooming and rusting might be distinguished from their decoy counterparts. And his hint is that in the former but not the latter cases, behaviour is "attributable" to the subject in question. But taken on its own this hint does not provide much insight. For the most natural interpretation of it is that, in the former but not the latter cases, the pertinent subject is *doing* something. And while it is true that, in moving its own legs, a spider is *doing* something, it seems equally true that our manipulated spider is doing something: It must at least be yielding to the strings that force its legs to move.

It might now seem tempting to reply that self-movement is a doing of a different kind than manipulation-yielding, and to suggest that by circumscribing the former, their contrasting natures might be elucidated. And it is true that these are different kinds of doing, discernible in all manner of ways: A spider's own movements have a typical causal history (inside the spider's body), typical outcomes (pursuing the spider's goals), and occur in typical settings (ones not involving strings pulling the spider's legs), and these things are not typical of

²¹Setiya (2011b, p. 154, n. 18) acknowledges complementary suggestions by Thompson (2008) and Alvarez and Hyman (1998).

the considered sort of spider manipulation. But this is far from showing that by describing a combination of such symptoms we could fashion a clear and general criterion, including all sorts of self-movement, and excluding all possible odd ways of manipulating a spider's limbs.²² And this latter suggestion seems much less credible. The mistake in it seems, by the way, analogous to the mistake of supposing that those doings which are an agent's actions are to be distinguished by some clear-cut overt mark, in the way that a facial expression's being sad might be.²³

Neither will it do, I think, to try to answer any of our four questions by appeal to some notion of being active, as opposed to being passive. Alvarez and Hyman (1998, pp. 243-245) may be right when they suggest that non-deviant mountain climbers, like rusting pieces of iron, are *agents* of the pertinent outcomes, whereas their mimicking counterparts are not. But if this suggestion is to explain these contrasts, then of course something needs to be said about this sense of agency. This notion must go beyond the mere notion of doing something, since every considered subject is doing *something* in every considered case. But as before, the notion cannot simply be defined by trying to characterise some overt mark which distinguishes these "agential doings" from "mere doings". And its proponent cannot fruitfully fall back on the Frankfurtian rhetoric of attributability, saying that a subject's being active in producing an outcome is a matter

²²Similarly, we may say that for a piece of iron to *rust* (as opposed to meticulously having a thin outer layer replaced with rust particles) is for a specific chemical process, presumably involving iron and oxygen, to occur in that piece of iron. Similarly, we may say that for a flower to *bloom* is for some process involving saps and light and I don't know what to drive its petals apart in a way that presumably runs along the lines of its recent evolutionary predecessors. But again these observations do not seem to help us make up a clear criterion which answers to the Frankfurtian challenge.

²³Here I am alluding to Anscombe's (1963, p. 30) attempt to undermine the suggestion that acting is to be distinguished by some characteristic, in the way that a sad expression might be. The way in which this suggestion arises in Anscombe's discussion is of course different from the way in which it arises in mine, and Anscombe has a sophisticated but somewhat obscure strategy for undermining this suggestion, which I do not pursue here. All I am saying is that the suggestion Anscombe wants to undermine is analogous to the suggestion that a spider's self-movement is constituted by some very general notion of movement plus some further characteristic of such movement.

of that outcome being attributable to this subject, since this is the notion we started out needing to elucidate.

In their talk of being active, Alvarez and Hyman do however make multiple appeals to the notion of a power, and the associated notion of exercising a power. Hyman has later suggested that the deviant causal chain problem “is not a problem about desires in particular, [but] about dispositions and powers in general” (Hyman, 2014, p. 18). The suggestion seems to be that we can identify the elusive notion of something’s being *active* in producing an outcome — and thereby, I think, the Frankfurtian notion of this happening being *attributable* to this subject — by saying that the subject in question possesses a disposition, and that the outcome in question is the exercise of that disposition. This line is, I think, promising. But Hyman does not make explicit what this notion of exercising a disposition comes to.

In making sense of this suggestion, we face an immediate hurdle in that there seem to be two main ways of developing it. On one understanding, the idea would be that the simple generic notion of exercising a disposition applies in each of the four cases we’ve been concerned with, and none of the corresponding mimicking cases, thereby neatly explaining the elusive relationship which exists between an agent and their action but which is not exemplified in deviant outcomes. On another understanding, the generic notion of exercising a disposition is not enough to explain this relationship. For on this second understanding, we need, for the case of action as for each other case raised, to appeal to the idea that what happens is an exercise of a particular *kind* of disposition. I want to pursue this latter approach. But first I want to say what is unsatisfactory about the first.

2.3.1 Exercising a Disposition

Someone might think that, because the problem of distinguishing acting from deviantly doing is an instance of a more general problem of distinguishing *exercising a disposition* from *doing something that mimics that*, all that is needed to solve the former problem is some general conception of *exercising a disposition*. Such thoughts seem to be suggested by some passages in the writings of Mele and Thalberg,²⁴ and may be reinforced by reading a brief pronouncement on part of Setiya (2011b).²⁵ The sort of theory resulting from such a thought would

²⁴Given the provisional way in which the points are made, it is time-consuming and potentially unproductive to pin the below form of account to these authors. But here I make the beginnings of an attempt at such an interpretation. Mele, drawing some inspiration from Thalberg (1984), proclaims that “causal accounts of what it is for an action to be intentional (unlike causal accounts of what it is for an event to be an action) cannot be falsified by waywardly caused nonactions” (Mele, 1992, p. 224, n. 4). Mele’s suggestion is clearly that there is some one general notion which precludes deviance but which may be innocently presupposed in an account of acting (or, in his corresponding terms, acting intentionally). It is less clear what Mele means by “action”. But given Thalberg’s sparse comments on the matter, it is at least reasonable to think that it is some notion of what happens being *due* to the acting subject — still in a sense that a theory of intentional action can comfortably take for granted:

[T]he causal theory we are talking about presupposes that we can at least in practice usually differentiate, among episodes of human behavior generically speaking, those special cases of behavior which rank as action. When *someone’s body moves*, we are normally able to decide whether *they moved their body*. (Thalberg, 1984, p. 249)

This notion of agent-due behaviour might, if only because it is so vague, be identified with the notion of being an exercise of a subject’s disposition. If it is not identified with that notion, I do not understand which notion it might be.

²⁵Setiya there sketches a proposal which seems structurally similar to Mele’s and Thalberg’s. This time, feeling reasonably confident about the intent of Setiya’s formulations, I have opted to make some terminological adjustments to avoid confusion:

[O]ur question is not whether [...] we can say what it is for the flower to open its petals (rather than having them opened by something else or by accident) in terms that do not presuppose its doing anything. The point is that *doing something* is a completely general topic in the metaphysics of agency whose generality is obscured by the restriction of ‘agency’ to rational agents. Call it ‘agency’ or not, there is such a thing as the exercise of a power or capacity by an object, inanimate or otherwise, about which we can ask: can this be explained in other terms? Someone who answers no, and therefore helps himself to the idea of [exercising a disposition], may nonetheless insist on a reductive account of what it is to [act]. This would be a causal-psychological theory of [action] without a causal theory of [exercising a disposition]; and it is the kind of the causal theory defended here. (Setiya, 2011b, p. 137)

But Setiya does not seem to think that this meaningfully contributes to solving the problem of deviant causal chains. He rather seems to think that it contributes to solving Velleman’s version of the problem of disappearing agents. However, even on this count I find the usefulness of this two-factor view dubious. As I have argued in footnote 9 on page 20, Velleman’s question can be disambiguated into two questions. Neither seems very well answered by the present

have the following shape:

Standard Causalism + Dispositionalism N acts in *Aing* *just in case* N has a desire that represents N as *Aing* (in a sense such that N can *A* and represent N as *Aing* without N's acting in *Aing*) and this desire causes N to *A* (in a sense such that this desire could cause N to *A* without N acting in *Aing*) and N's *Aing* is an exercise of a disposition on part of N

The problem with this proposal is that, because the invoked notion of exercising a disposition is generic, the account makes no real headway with the problem of deviant causal chains. A subject can apparently exercise some disposition to *A*, and conform to standard causalism's conditions with respect to their *Aing*, and still not act in *Aing*. Davidson's climber example seems to show this with just slight elaboration, since it conforms to standard causalism's conditions, and since, in addition to this, letting go of a rope because of a nervously loose grip apparently amounts to exercising some disposition to let go of such ropes.²⁶

suggestion. If Velleman's question is read as Nagel's concern about identifying an agent within a flux of events, bringing in the agent and their capacities might be a promising start for answering it, but it is not clear how this suggestion, on its own, can explain in what sense agents are the doers of their *actions* (though perhaps it can explain some more general sense in which agents are doers). If Velleman's question is supposed to concern "that which distinguishes human action from other animal behaviour" (Velleman, 1992, p. 462), then it is not clear how Setiya's introduction of the generic idea of exercising a disposition answers it, since humans and animals *alike* seem to exercise dispositions in acting.

²⁶Analogous remarks seem applicable to the other considered cases. The generic idea of exercising a disposition is not much use, on its own, in saying when a spider moves its own legs, since even if someone manipulates a spider's legs with strings, the spider's resulting movements may be described, in a contrived but basically kosher manner of speaking, as an exercise of *some* disposition on part of the spider (like the tensility of its legs). Lewis (1997, p. 155) apparently would take the contrary view that, in such a case, it is not the spider's dispositions that are manifest, because the spider's "intrinsic causal bases" are not involved. Lewis does not give an argument for this restriction, and it is not so clear to me what an intrinsic causal base is supposed to be. Below I ignore Lewis's view on the matter, but hold out some hope that Lewis's dispositions might be identified simply as a subset of my dispositions — as those which, for some more or less principled reason, are judged to be more deeply grounded in the features of their subjects.

2.3.2 Exercising a Kind of Disposition

Our first thought was that acting, like blooming, involves exercising a disposition. But our second realisation was that the mimicking counterparts of acting and blooming involve exercising dispositions as well. Hence it can seem that the first thought does not contribute to separating acting and blooming from their mimicking counterparts. But there remains a possible way of drawing out the difference between the first pair of notions and the second pair, which still makes a fundamental appeal to the notion of exercising a disposition.

For having noted that there is some sense in which, say, the nervous mountain climber who deviantly lets go of a rope exercises a disposition (and even a disposition *to let go*), it still seems irresistible to qualify this, and say that, although the nervous mountain climber does exercise some disposition (subserved in part by their nervousness), they do not exercise the *kind* of disposition which is exercised in action. Similarly, it might be that a manipulated flower exercises some of its propensities, but it does not seem to exercise the same sorts of propensities that are involved in *blooming*. And similar observations seem to apply for the other considered cases.

So let's conjecture that, although we cannot understand what it is to act by pairing the generic notion of exercising a disposition with further separate notions, we *can* understand it as exercising a disposition of a kind. The conjecture is that, for some kind of disposition, exercising it *is* acting. Analogously, suppose that blooming, moving and rusting are exercises, respectively, of three further kinds of disposition. In a way, this gives a rather direct answer to Frankfurt's problem of telling

the difference between what goes on when a spider moves its legs in making its way along the ground, and what goes on when its legs move in similar patterns and with similar effect because they

are manipulated by a boy who has managed to tie strings to them.

(Frankfurt, 1978, p. 162)

For it answers this question, not by distinguishing the overt characteristics of the two kinds of leg movement, but by saying that what goes on when a spider moves its legs is an exercise of a particular kind of disposition on part of the spider — a disposition which is not exercised (and perhaps not even present) in the case where someone manipulates the spider’s legs with strings. Similarly, to Frankfurt’s question of what the difference is “between the sort of event that occurs when a person raises his arm and the sort that occurs when his arm goes up without his raising it” (Frankfurt, 1978, p. 162), we may answer that the first kind of raising but not the second is an exercise of some pertinent kind of disposition on the part of this person.

But as direct as the above answers to Frankfurt’s questions might be, they are also incomplete. Completing them requires expressing, for each case of the above sort, *which kind* of disposition it is that is exercised in one case and not in the other. That requires, in turn, a serviceable conception of how to divide dispositions into kinds, and equally a serviceable conception of what it means for those dispositions we’re concerned with to be exercised. But there is, at least, a readily available way to divide dispositions into kinds, which may with slight elaboration be supplemented with a conception of what it is to exercise a disposition of some kind. To see this, start with a widely held view of the nature of dispositions, largely due to Ryle (1949). The view is as follows:

Vague Conditionalism For N to be disposed to A in condition C is for N to be the type of subject that, if condition C obtains, N will A.

There are, of course, many debates about the nature of dispositions. But most of the disagreements embodied in these debates should not, I think, be construed as concerning the truth of *Vague Conditionalism*. Theorists working on

dispositions tend to take it for granted, as a starting point of their enquiries, that “[s]tatements ascribing causal dispositions or powers are *somehow* linked to (strict or strong) conditional statements” (Martin, 1994, p. 2), that (in other words) “disposition ascriptions have something to do with conditionals” (Fara, 2005, p. 46), or that “there is some connection between dispositional properties and counterfactual conditionals” (Choi, 2008, p. 795). This is not enough to show that all endorse *Vague Conditionalism*. But I think it can be shown, at least, that one especially prominent and deep disagreement about dispositions can be understood as a debate between parties who accept *Vague Conditionalism*:

Some, like Choi (2008), account for a thing’s dispositions by appeal to straightforward conditional facts about what the thing would do if a given condition obtained. So they seem to think that dispositions correspond to exceptionless generalisations concerning what subjects would do in some possible conditions. The viability of Choi’s theory is not my present concern. My point is only that a theory like Choi’s is not a rejection of *Vague Conditionalism*, but just a particular way of understanding that thesis, which views its apparent conditional as an actual and straightforward conditional.

One apparently compelling ground for questioning Choi’s kind of view is that a thing can have a disposition, and that the condition associated with the conditional can obtain, although — perhaps because there are preventing circumstances — the subject with the disposition remains inert (Fara, 2005, p. 61). But having claimed that straight conditionals won’t do, such a theorist will be likely to introduce something which is very much *like* a conditional *except* in allowing for such cases of inertia. Someone might, for example, say that having a disposition is being such as to do a pertinent thing when some condition obtains *unless there are abnormal or interfering circumstances*. Or it might be

suggested that having a disposition is being such that, *in most possible cases* where a condition obtains, the subject with the disposition will go on to do a given thing. Each such proposal can be seen as a way of endorsing *Vague Conditionalism*, while trying to improve on Choi's understanding of the force of the apparent conditional that figures in this thesis.

Fara's own approach is more sophisticated. It proceeds by introducing habituals, such as are employed when we say that a certain bird migrates during summer or that someone eats two portions when there's dinner (Fara, 2005, p. 66). Someone else might have stopped there, simply identifying the fact that someone has a particular kind of disposition with the idea that some habitual is true of them. Again, this would be a way to conform to *Vague Conditionalism*, elaborating on it in yet another way. But Fara does not stop with such an account. To him, a subject that has a disposition must not only make such a habitual true, but must make it true in virtue of possessing some *intrinsic property* (Fara, 2005, pp. 69-70). Whether or not the resulting account is deemed plausible, it seems that it, like the previous suggestions, is a way of providing sense to, rather than a way of refuting, *Vague Conditionalism*, this time viewing its apparent conditional as constituted by a habitual claim which sets certain extra demands on the subject which it is about.

Hopefully this brief excursus is enough to provide some assurance that *Vague Conditionalism* is an uncontroversial thesis about dispositions. In any case, assuming this thesis affords us with information about what it is to exercise a kind of disposition. For assume that a Rylean, conditional-esque statement, is true of a subject. Perhaps it is true of a piece of glass that, if struck, it will shatter. We may then say that this piece of glass is disposed to break when struck, and thereby that it possesses a certain kind of fragile disposition. Now what is it for this piece of glass to exercise this kind of fragile disposition?

Minimally, the antecedent of the Rylean conditional must obtain (the piece of glass must, in fact, be subjected to a strike, or — depending on just how the details of the account are worked out — have been struck), and the consequent must of course be true as well (the piece of glass must be breaking).²⁷

Since we have proposed to define each of the four kinds of doing we are presently concerned with *as* exercises of pertinent dispositions, this means that, if something is doing any of *these* things, then a pertinent Rylean conditional (exceptionless or not), must be true of the subject doing it, and, further, the antecedent of that conditional must obtain. So if a spider is moving its own limbs, then such a conditional must be true of the spider (perhaps the conditional is that the spider is such that if it perceives such-and-such event with relevance to its spider interests, then it will move its limbs), and the antecedent of that conditional must be true as well (things must be such that the spider does perceive an event of that type).

This defends the idea that spider movement is different from spider manipulation, since now we can see that the former but not the latter requires a Rylean conditional to be true of the spider (a dead spider, not such that it would move if the perceptual condition obtained, could of course still be manipulated with strings as a live one might be), and that the former but not the latter kind of movement requires that the condition associated with this disposition obtains (a spider which did not register its web as being shaken could still be made to move as if in response to such shaking, but could not exercise the pertinent disposition).

And we may extend this approach by introducing similar Rylean conditionals

²⁷A difficult issue remains over whether, if a piece of glass is disposed to break when struck, it could be struck and break without exercising the pertinent disposition. If that is possible, then exercising that disposition could not be given a straightforward definition in terms of possessing it, being struck, and breaking. But this need not concern us here, since we are not trying to give an account of what it is for some disposition to be exercised, but rather assuming the notion of exercising a kind of disposition, and examining what this notion entails.

in the hope of capturing each further kind of exercise, thereby distinguishing each further kind of process from its decoy counterpart:

Movement For a spider to move its own legs is for it to exercise a disposition to move its legs when an associated condition obtains (plausibly that the spider perceives such-and-such)

Blooming For a flower to bloom is for it to exercise a disposition to bloom when an associated condition obtains (plausibly that there is some light for some length of time)

Rusting For a piece of iron to rust is for it to exercise a disposition to rust when an associated condition obtains (plausibly that the iron encounters oxygen)

The parentheses contain suggestions for how to state the condition associated with each kind of disposition. There is room for debate about how to sharpen or adjust these. Here I am more trying to illustrate an approach to answering Frankfurtian questions than provide a very full and fine-grained answer to each Frankfurtian question that might be raised. (After all, this is not a thesis about what it is, for example, for something to be rusting.) The general approach is to identify each kind of process as the exercise of a kind of disposition, and each kind of disposition by appeal to an associated condition, thereby exposing how, because mimicking counterparts of these kinds of doing do not require these dispositions, nor that the associated conditions obtain, they are different in kind from these doings.

Assume, now, that an answer along these lines could be given for the case of action, so as to finally show what distinguishes someone who is acting in letting go of a rope from someone who deviantly happens to let go of it, and also someone who is acting in raising their arm from someone whose arm is

raised by some external force, and any other case of acting from any other such mimicking case which someone might come up with. Acting must hence be understood as the consequent of a conditional of the kind which describes a disposition. The consequent of the conditional must simply be that someone is acting. But what condition could plausibly figure as the antecedent of such a conditional? Inspired by Anscombe, but also, in a way, by the Davidsonian tradition, I propose the condition that the subject who is acting believes they are acting. Hence the proposal takes on the following shape:

Acting For N to act in Aing is for N to exercise a disposition to act in Aing in conditions where N believes that N is acting in Aing

Somewhat in line, for example, with Coope (2007), I propose to call the kind of disposition named by *Acting* a desire. In one swoop this will defend the attractive thought that someone lacking in dispositions to do something could hardly be said to want to do it, and the nearly undeniable thought that a desire can only give rise to an action in the presence of a belief that represents the wanted course of action (Hyman, 2014, p. 4). I think that calling the presently described kind of disposition a desire is useful in more ways, both for making the dispositionalist account seem more familiar than it might otherwise, and for underpinning more philosophical instincts about the nature of desire and the way it comes together with a belief to produce an action.²⁸

Aside from any further suspicions one might have about the truth or usefulness of *Acting*, it is quite clear that it solves the problem of deviant causal chains when this problem is construed as a Frankfurtian problem of distinguishing acting from happening to do something which is similar to acting, but is

²⁸I suspect, for example, that the issue of whether and how weakness of will is possible can now be fruitfully viewed as a question of whether a desire's condition can obtain without the desire being exercised (without, that is, the agent acting on it), making this into a specific instance of a more general problem of whether and how it is possible for any disposition's condition to obtain without the disposition being exercised (something known as "masking" in the literature on dispositions).

not that. For now we may say not only that acting is exercising a particular kind of disposition, unlike doing something which mimics acting, but that acting requires a subject to want to act, and also to believe that they are acting, as hardly seems necessary in the mimicking cases. An unconscious subject, lacking a desire to let go of a rope, could still let go of a rope in the way that the nervous mountain climber deviantly does, but could not act in doing it. A conscious subject, wanting to perform an action of letting go of a rope, but unaware of doing so, could not act in letting go of a rope, but could let go of it in the deviant way.

This solution to the Frankfurtian problem is just another instance of the form of solution sketched for the other cases. The general form of such solutions is to say that, because a certain kind of doing only counts as happening when it is an exercise of a certain kind of disposition, and therefore only where the condition that triggers this disposition is satisfied, occurrences which do not place such requirements are not of this kind. But each concrete proposal for how to individuate a disposition might be questioned. For example, does a spider that is moving its own legs really have to have a disposition the triggering condition of which is somehow perceptual? I will ignore such questions as they pertain to spiders and the rest. But in the next chapter I will argue, in support of *Acting*, that the present kind of disposition, and the present kind of condition, are constitutive of action. In the course of so arguing I will also endeavour to clarify the belief condition which gives sense to *Acting*.

Chapter 3

Doing What's Believed

3.1 Introduction

The starting point of this chapter is the following pair of thoughts: Acting is exercising a desire, and a desire is a disposition to do what practical belief represents in the presence of such a belief. It argues that these thoughts can only be defended by adopting a third: Whatever a practical belief represents its subject as doing, it represents this subject as *acting* in doing it. This third thought comes with a methodological moral: Because practical belief figures in any satisfactory account of action, and because action figures as the object of practical belief, there can be no account of action which is not circular.

We then need to say something about the object of practical belief that goes beyond the bare claim that it represents its subject as acting in doing what they're doing. My method for drawing out further commitments about the object of practical belief is considering various objections to the bare claim, and showing how it needs to be understood for these objections to miss their mark. This results in the three claims that practical belief is *imperfective*, in

that it represents its subject's action as ongoing, and *causal*, in that it represents an ongoing action as a matter of doing something which causally contributes to it, and *part-homogeneous*, in that it represents these causally contributing happenings as further actions.

3.2 Practical Belief

Our attempt to understand action through desire, desire as a special kind of disposition, and this kind of disposition through its association with a belief about what's done, has led us to the claim that whenever an agent acts in doing something, they are aware of doing that (Anscombe, 1963, p. 25). In order for this belief condition to make good on this dispositionalist account, the condition needs to be developed in such a way that no one can act without satisfying the condition, and so that no one can satisfy the condition, and exercise the associated disposition to do what is represented by the belief that figures in it, without acting. What sort of belief could have this dual role?

I propose that, to complete the dispositionalist account, it is both required and helpful to introduce, in the role of a desire's condition, a belief that represents its subject is acting in *A*ing. This condition clearly provides for the need to make exercising a desire without acting impossible. For of course no one can verify a representation of themselves as acting without acting, and hence no one can exercise a disposition to verify such a representation in its presence without acting. This is how the account maintains the advantage of making counter-examples in the style of deviant causal chain cases impossible. (Unwitting or deviant watering of flowers can't correspond to a representation of someone as *acting* in watering flowers, and hence can't amount to exercising the kind of disposition at issue)

But is it really true that, in acting, a subject must believe they're acting?

After all, it has appeared false to a number of philosophers that a subject must even believe they're *doing* whatever they're acting in doing. And at any rate, just what does it mean to say that, to act in doing something, a subject must believe they're acting in doing it? This pair of questions can be combined into a challenge: What is it to believe someone is acting in doing something, such that anyone who acts believes this about themselves? This is the question I want to answer in this chapter. I will do it by considering several objections to the claim that a belief about acting is a precondition for acting, all apparently built on top of counterexamples, and by showing a way to think of the object of a practical belief which allows the thesis to avoid the counterexamples.

Before going on to do this, I want to say why it is necessary. Does the emerging account require introducing the idea, not only that some understanding of what one is doing is necessary for acting in doing it, but also that, if one is acting in doing something, one must believe this very fact? I think the answer to this question must be positive. It is true that some other kind of belief, which doesn't represent its subject as acting, might have the feature of being omnipresent in action. But because it is always possible to have such a belief and verify it without acting,¹ it will always be possible to exercise a disposition to satisfy such a belief in its presence without acting. For that reason we could not rework our dispositionalist account to be founded on any such weaker condition.

Sensing that, in introducing a belief specifically about acting, we have abandoned some implicit methodological constraint, someone could now pointedly ask why we should not go further, by doing without a belief condition and bak-

¹It should be obvious from previous lines of investigation that we can't reform the weaker belief condition in such a way that satisfaction of such a belief will be sufficient for action, since, for example, there's no *way* of Aing that seems so much as necessary for acting in Aing (except the "way" of Aing which is acting in Aing), and since seemingly no kind of causal route from a belief about Aing to Aing which will guarantee acting (since attempts to formulate a sufficient condition by talking about causal relations (in an action-independent sense) are either false or versions of the mysticism-exemplifying claim that acting is distinguished by an unknown feature that is specific to acting).

ing the relevant constraints into our desire condition, or by taking the reverse approach and leaving desire out of the picture. Taking the first line, one might say that a desire is a kind of disposition to *act*, whatever the condition is in which a desire manifests.² This claim, bare as it is, seems true to me, and is a consequence of the account I favour. But on its own it gives us no hold of which kind of disposition it is that manifests in action. The account I favour allows us to say that it is a kind of disposition which is *exercised* when its subject *believes* they're acting.

Suggesting, instead, the second line, someone could say that to act in Aing is to think one is acting in Aing in a case where this belief is true. As is quite obvious, this is also a consequence of my account. But without the introduction of desire, this account offers no answer to the question of what it is for a subject to be such as to verify such a belief. My preferred account is in a position to give at least this answer: For such a belief to be true is for the agent to be *disposed* to do what the belief represents in conditions where the agent has that belief (hence wanting to do it), and for this disposition to be *exercised* (hence doing what the agent wants to do).

My account elaborates on these simpler accounts but shares their circularity. If it is true, it seems that no further elaboration will extinguish the circularity. For it does not seem possible to get hold of the relevant kind of belief without specifying its object, which reintroduces acting into the account, or the relevant kind of disposition without specifying what constitutes exercising it, which also reintroduces acting.³ Is this, in itself, an objectionable feature of the account?

²Of course on any concrete version of this theory, the condition “parameter” would have to be filled in. Perhaps a proponent would say that desire is a disposition to act in a conditions where one is in some specific kind of neural state, or in conditions in which one has some belief about the good of doing something, or in which one would enjoy doing it. Such accounts seem to have counterexamples — are robots, the bad, and the stoic somehow banned from acting? A deeper but more elusive issue with these accounts is that, rather than explain what is involved in acting, they seem to just introduce elements alongside every action with no clearly elucidated, nor clearly elucidating, conceptual connection to it.

³This point seems to be Wilson's (1989, p. 275).

It seems rash to object to the account on grounds of some general prohibition against circularity. But is there a good objection that targets the more specific circularity of accounting for acting by introducing a belief about acting?

Some wariness about this specific kind of circularity seems to tacitly constrain many theorists who write on action, since very few openly endorse or consider accounts with this shape. One of the few who does openly consider a proposal like this is Peacocke:

Perhaps then it may be said that there can never be a *perfect* match with a given intention to φ in these deviant cases, because the intention to φ is the intention that one should φ as a result of one's possession of this very intention via a nondeviant chain. (Peacocke, 1979, p. 57)

But Peacocke immediately objects to this on the following alluring but unclear grounds:⁴

In these examples the reason that the token actions do not match such intentions is simply that they are produced by a deviant chain. The suggestion gives no clue about what it *is* for a chain to be deviant. (Peacocke, 1979, p. 57)

Reformulating Peacocke's thoughts, we may express his offhand suggestion as an attempt to exclude deviant causal chain cases by saying that to act, it is necessary that one veridically represents oneself as *acting* (and not as deviantly happening to do something), and his subsequent remarks as an objection to the

⁴Before this, Peacocke presents a different complaint, which I ignore in the main text. It is that "such a view wrongly involves ascribing to anyone with intentions use of a distinction between deviant and nondeviant chains of which he may never have dreamed" (Peacocke, 1979, p. 57). Of course whether or not an agent may have dreamed of such a distinction depends on what the distinction amounts to. If, as Peacocke thinks, knowing that difference is a matter of knowing what differential explanation is, there is ample reason to think agents don't know it. But if, as I think, the only sound distinction to know of in this vicinity is that between acting and doing something which is not acting, but which is perhaps similar to it, then there seems to be no difficulty about saying every agent knows it.

effect that, because acting figures in this condition (as something represented) the account does not tell us what it *is* to act.⁵

The objection could be read in two ways. On one reading, the objection is that this account of acting in terms of a belief about acting could, but does not say enough to characterise the object of that belief. But then this objection seems to just be an especially pointed way of asking the question which this chapter is devoted to answering: What could it be to believe that one is acting, such that such a belief is a necessary constituent of the fact that someone is acting? On another reading, which has not yet been considered, the objection is that if there's an account of acting which introduces a psychological state specifically about acting, then this account *can't* characterise the object of that state, or at least can't do it fully. And of course this would mean that this account couldn't say what acting is, or couldn't do it fully. If it is good, this objection is fatal to the present account.

Although there is something compelling about the second objection, which concerns the informativeness of the present form of account, I believe that a discussion of the first kind of objection will put us in a better position to answer it. Grasping a determinate account of acting which introduces a belief about acting will, that is, be helpful in staving off the concern that no such account could be informative. Hence in this chapter I bracket the ponderous concern about informativeness, and consider only the first kind of concern: What might the object of a practical belief be, such that someone who is acting could not fail to have such a belief? My method for answering this question will be to consider

⁵Wilson similarly objects to such ideas by saying that “[e]ven after we have been told that a state of intending refers to itself, we still do not know *what* state[...] is purportedly referred to” (Wilson, 1989, p. 279). This objection to what Wilson calls “self-referentially” is tied up with the assumption that beliefs are tokens of language of thought sentences. Setiya (2007, p. 46, n. 38) rejects that extra assumption, and suggests that this helps to avoid Wilson’s objection. But at its core, Wilson’s objection seems to be the same as Peacocke’s, and hence rejecting Wilson’s largely Fodorian conception of psychology seems to be an unsatisfying way to respond to it.

some purported ways of showing that a practical belief has no necessary place in action, and to refine the belief condition in such a way that these attempts miss their mark.

3.3 Practical Belief is Imperfective

In this section I want to argue that, although someone who is acting in doing something must believe they are acting in doing it, they need not therefore believe they will ultimately be successful in doing it. The discussion can be started by considering one aspect of Anscombe's work on the matter.

Anscombe strives to define acting as the object of an attitude she calls "practical knowledge".⁶ She discovers a dilemma: On the one hand, practical knowledge is supposed to be knowledge of what a subject does — of something that happens (Anscombe, 1963, pp. 52-53). On the other hand, reflection on failed actions, such as one where someone is ordering the building of a house but their orders are swiftly disobeyed — suggests that someone can have such knowledge of their action even though nothing much actually happens (Anscombe, 1963, pp. 55-56).⁷ I paraphrase the resulting dilemma as follows:⁸ How can the atti-

⁶In the subsequent text I do not draw very much on Anscombe's account of practical knowledge, partly because when this account is clear, it tends to be negative and roundabout, whereas when it is positive, it tends to be obscure. For example, the main negative claim, which seems to have attracted the most attention, is that practical knowledge is not based on prior evidence (Anscombe (1963, pp. 88-89). This does not tell us very much about how we should conceive of practical knowledge, or what it is based on, or how it is based on it — questions that have become focal to that just-mentioned attention. The main positive claim seems to be that practical knowledge is "the cause of what it understands" (Anscombe, 1963, p. 87) — which (similarly to the account which starts this chapter off) does not seem to tell us very much about what it is that is understood and caused by such knowledge.

⁷A minor remark on exegesis: The way Anscombe poses the problem can make it seem as if she thinks that the problem stems from the simple view of knowledge as factive, and that her solution is to suggest that the kind of practical knowledge that a subject has of what they are doing is non-factive. But she never clearly defends such claims in her treatment of practical knowledge, so this reading ultimately seems unwarranted.

⁸I bracket Anscombe's seeming and connected thesis that where the object of practical knowledge *does* happen, this practical knowledge does not require observation of it. Contrary to a common way of approaching Anscombe's view of practical knowledge, I will focus, in this chapter, more on the shape of the objects of practical knowledge than on questions concerning how such knowledge is (non-observationally? non-inferentially?) *based* on its

tude involved in acting be one of knowledge if it does not require the reality of its object (Anscombe, 1963, p. 82)?

The force of the dilemma is readily felt: Isn't there a vivid sense in which someone always knows the action they're doing, even if unbeknownst to them things are going very badly? How then can their knowledge survive such total but unknown misfortune in bringing it about? But it is not clear just what its incompatible lemmas are supposed to be. McDowell, for example, is "not sure what to make" of the dilemma Anscombe seems to express (2010, p. 430), and Haddock (2011, pp. 167-168) gives voice to some uncertainty about which parts of Anscombe's various interjections are attributable to Anscombe's official self, as opposed to being spoken by an imagined interlocutor.

I want to start by offering a quite straightforward interpretation of Anscombe's concern, on which it roughly amounts to the following question: How can acting require an attitude that accurately represents what one is doing if one might ultimately fail to get it done? It is true that the present interpretation hardly exhausts Anscombe's anxieties on the matter of practical knowledge. But it is possible to attribute to Anscombe a desire to resolve the question just stated. And in any case, stating and resolving the above straightforward question will force us to appreciate the imperfective nature of practical belief, thus helpfully clarifying a more elusive and difficult version of Anscombe's concerns, developed recently by Setiya.

To more clearly appreciate the force of the initial, simple question, consider first the below pair of theses:⁹

Practical Belief: If N is acting in Aing, N has a true belief that N is Aing

object (Haddock, 2011), and on related questions of how basing knowledge on something not-necessarily-currently-real can be *rational* (as opposed to a form of wishful thinking) (Paul, 2009b,a; Setiya, 2008b, 2009, 2011a).

⁹I will speak of true belief where Anscombe speaks of knowledge — the difference between these attitudes will not, I think, matter to Anscombe's dilemma, nor to its solution.

Failure-tolerance: N can act in Aing even if N will not end up having A-ed

As can be read straight off this pair of theses, they are not inconsistent unless “N is Aing” is silently interpreted in such a way that it entails “N will end up having A-ed”. Only then does one of the theses say that acting requires true belief in a kind of fact which, according to the other thesis, need not exist where someone acts. But there is no apparent reason to silently interpret things in this way, as, on the face of things, believing that one is doing something is different from believing that one will end up having done it. If, after realising this, we slightly adjust *Practical Belief* to try to capture the intuitively felt dilemma, we will of course generate what looks like a real inconsistency:

Practical Belief*: If N is acting in Aing, N has a true belief that N will end up having A-ed

Failure-tolerance: N can act in Aing even if N will not end up having A-ed

Practical Belief is consistent with *Failure-tolerance*, but *Practical Belief** seemingly must be inconsistent with it. A straightforward reconciliation of this latter pair of theses hence seems impossible. But someone who, for some reason or other, favours *Practical Belief** might try to resolve the tension simply by rejecting *Failure-tolerance*. As I’ll soon argue, this is misguided, since simple observations about the possibility of failing to complete an action undermine *Practical Belief**, and give at least provisional support to *Failure-tolerance*. This will leave those who are attracted to *Practical Belief** the option of modifying this thesis to accommodate *Failure-tolerance*. Again this will be seen to be a misguided move. All this will leave us with *Practical Belief* as the only plausible way of defending the omnipresence of practical belief in action.

Rejecting Failure-tolerance in Favour of Practical Belief*

Straightforward cases, which do not involve especially unfortunate epistemic circumstances, can seem to undermine *Practical Belief**, and provide some support to *Failure-tolerance*. For example, we may imagine that someone is building a house in the ordinary way, with some sensory or testimonial feedback. Such a person may be acting in building a house, although they have not yet built a relevant house. There may not even be a foundation in place. And of course it is then possible that the building will end in failure. Plans can change, and foundations can be sucked into swampy grounds, ruining builders. So for this particular case, at least, something along the lines of *Failure-tolerance* must be true. Such common-sensical points, noticed and insisted on by a number of philosophers (Falvey, 2000; Setiya, 2011a; Thompson, 2008), seem to falsify *Practical Belief**, and seem to lend at least partial support to *Failure-tolerance*.

Though the interpretive issue is hardly straightforward, it can seem that some of Anscombe's own examples in her discussions of practical knowledge are intended to defend *Practical Belief** from such points. In one of these examples, someone is aware of having erected a house by giving orders, but hasn't gotten any feedback about how the building is going, and in fact unbeknownst to them their orders have been disobeyed, and nothing has gotten done. Here Anscombe can seem to claim that there is practical knowledge of the fact that a house has been built, although at the same time, no house has been built (Anscombe, 1963, p. 82). This claim can seem designed to make way for *Practical Belief** even if something like *Failure-tolerance* is true, at least for some cases. For if someone can have practical knowledge of what has happened without it having happened, then it is not far off to add that they can know of an action they *will* have performed without things being such that they will have performed it. But the resulting, "non-factive" notion of knowledge or true belief seems incoherent

to me,¹⁰ and thus I find it hard to believe that Anscombe is, in fact, proposing some such conception of practical knowledge, still less that she wants to define action as the object of that kind of knowledge.

Rejecting *Practical Belief** in Favour of *Failure-tolerance*

But could some other similarly perfective attitude be definitive of action in the way that *Practical Belief** fails to be? As far as I can see, there are three possible axes along which to vary *Practical Belief** in the hopes of providing an account of the kind of understanding of what one is doing which defines action: First, one could try to think of a different *type* of perfective attitude, and introduce some version of the claim that a subject acts in Aing just in case they hope, or intend, or have some other kind of attitude directed towards successful completion of Aing, where whatever attitude is invoked is true or veridical or satisfied. Second, one can try to come up with qualifications to the *object* of practical belief, introducing something like the claim that so that acting in Aing is having a true belief that one will have A-ed *relative to some circumstances*, or at least that success is *possible*. Third, one can shift the object of practical belief, introducing something like the claim that acting in Aing requires having a true belief that one will end up having X-ed.

The first kind of proposal does not seem to improve on *Practical Belief**. For the simple argument against this thesis did not seem to turn on any of the possible differences between true belief, knowledge, hope, desperate hope, or vague hunch. As long as these attitudes are construed as veridically representing the fact that their subject will end up having done the action they are currently

¹⁰Anscombe herself notes the possible retort that “it is a funny sort of knowledge that was still knowledge even though what it was knowledge of was not the case!” (Anscombe, 1963, p. 82). I suppose the thought behind the retort is that knowledge cannot be properly understood if we try to construe it as a non-factive attitude. As far as I can see, subsequent parts of *Intention*, which are meant to articulate Anscombe’s notion of practical knowledge more fully, do not really address this kind of objection.

doing, then, as before, those cases which seem to conform to *Failure-tolerance* will threaten the omnipresence of such states in action.¹¹

The second kind of proposal does not fail in an immediate way, so let's consider it. The idea in it is that although true belief that one will end up having A-ed isn't necessary for acting in Aing (as is shown by the fact that any action can be truncated), true belief that one will end up having A-ed *if conditions are such-and-such* is thus necessary:

Practical Belief:** If N is acting in Aing, N has a true belief that N will end up having A-ed *provided* that p

But this thesis seems no more true than its predecessor. Firstly, it seems that an agent can act in doing something without being fully aware of what is needed for success. Someone might, after all, start doing something while pretty unsure about how to get it done. A possible response to this is that this uncertainty does not preclude belief. The suggestion would be that, even if there's uncertainty, someone who is doing an action must have some idea of what will result in success.

But consider the following more sophisticated objection: Whatever an agent thinks will be sufficient for success can turn out to not be sufficient ("oh, this lever isn't connected"), in which case a responsive agent can come to form new beliefs about what will be sufficient ("oh, there's another one"), all while continuously acting in securing the end "through an alteration in envisaged subordinate means" (Thompson, 2008, p. 105, n. 12). In such cases, where things go

¹¹If we relax the requirement for veracity, then is there a perfective kind of belief or hope that is necessary for acting in Aing? We may try the mere suggestion that "For N to act in Aing requires N to have a belief that N will end up having A-ed". If such a belief is such as to require the absence of very much doubt about whether its subject will go on to A, it is quite obviously falsified by cases like the one where Davidson imagines himself to try to be making ten carbon copies, since after all it seems true that in this case he does "not know, or believe with any confidence" that he is *succeeding* (Davidson, 2001a, p. 92). The question of whether some appropriately meek hope of future success is necessary for acting in Aing is more difficult (Setiya, 2008b, 2009). But as long as the attitude in question is not understood as necessarily veridical, it will be very hard to see how it could contribute to completing an account of what it is to act.

well though not according to initial plans, there seems to be no condition of the form “I will end up having A-ed, provided ...” which is necessarily *both* believed *and* true *throughout* someone’s acting in Aing. If there is no such condition, then there is no value assignment to “p” which can make *Practical Belief*** true.¹²

Someone could then say that while the agent might not necessarily have the right idea of which specific conditions will facilitate success, an acting subject must know that there are some such conditions. I think that proposal must take something like the following shape:

Practical Belief*:** If N is acting in Aing, N has a true belief that N *possibly* will end up having A-ed

It is hard to doubt that whenever someone acts in Aing, they know they *may* (in some weak sense of “may”) end up successful in their endeavour. And I will not try to come up with a wild counterexample to undermine this thesis (involving, I suppose, someone who is acting in doing something which is absolutely certain not to succeed, or someone who is acting in doing something which seems so difficult that they think it is totally impossible to carry it out).

Rather I want to object that, even if this attitude is necessary in action, it does not help provide for an account of what it is for a subject to act, since it is not specifically characteristic of action (the passive can have the same attitude) and does not in any core way seem to be related to action (the passive can satisfy the attitude merely by being such that they might have successfully done the action in question). As it turns out, a version of this leniency objection applies to all subsequent proposals too.

Consider the third and perhaps more interesting kind of revision, which shifts the object of practical belief, as below:

¹²Again, replacing “belief” with “hope”, “intention” or “knowledge” does not yield a different conclusion.

Practical Belief****: If N is acting in Aing, N has a true belief that N will end up having A-ed *or* X-ed (where X is different from A)

This seems to encapsulate Davidson's (2001a, p. 50) view of the matter, on which acting in ensuring the future welfare of one's children entails not knowledge that one will manage to do that (since the uncertain goal may even be beyond one's death), but rather knowledge that one will have taken pertinent legal steps with one's lawyer, or put one's signature on a piece of paper, or that one will have done something else, where presumably what is thereby done has some privileged connection to the action in question.¹³

*Practical Belief***** seems true to me. If someone is acting in doing something, like walking to school, they must apparently believably have done many further things (walked to school, or taken on shoes, moved their body, or ...). If Anscombe's builder is acting in building a house, they seemingly must be aware of some past successes along the way to building, such as just ordering people to build, or making a plan to build, or deciding to build, or whatever else might have been involved in the project thus far. These humdrum observations seem to vindicate the thesis in question. But the fact that past success is cheap is also what makes trouble for it, since, again, someone can know of past successes in doing things along the way to Aing even after having given up the whole project. Similarly to the last proposal, this one constructs a notion of practical belief that is too promiscuous to be definitive of action.

Having agreed that some of the above proposals give us a too strong belief

¹³For of course it is plausible that not just anything could figure as the "something else", or the X, mentioned above. I would say, maybe in agreement with Davidson, that it is some X taken as a means to Aing that necessarily figures in the thought of someone who is acting in Aing. Davidson would say, probably in disagreement with me, that acting in Aing requires knowledge of some Xing, where the Aing that happens is *identical* to the Xing known. This adds to *Practical Belief***** Davidson's recurring idea that actions are individuals that can come under many descriptions, and can be known under one but not another. But I don't think we need to concern ourselves with any such additions to *Practical Belief*****, since these don't seem to help build a defence to the upcoming difficulties about the temporal character of such knowledge of Xing.

condition, whereas others make it too weak, the present kind of theorist might suggest that the difficulties are due to a too binary conception of belief. They may then adopt a suggestion due to Setiya (2008b, p. 391), on which it is not all-out belief, but some comparatively high level of confidence in future success, which is definitive of action. (It is not clear that Setiya’s suggestion is made in the spirit of defending the present kind of theorist, who views practical belief as oriented towards future success, but this suggestion might be retrofitted to suit them.) Consider, then, the following thesis:

Practical Belief*****: If N is acting in Aing, N has a higher level of confidence that N will end up having A-ed than N would otherwise have, and N will end up having A-ed

As emerges in a debate between Setiya and Paul (2009b), making such a claim true requires careful curating of which type of situation to fit into its “otherwise”-clause (Setiya, 2009, p. 131). Should “otherwise” be taken to refer to what things were like right before the agent started acting, or to what things would have been like if everything were mostly similar except the agent weren’t acting? But rather than engage in this discussion, let’s provisionally grant that on some such careful interpretation, *Practical Belief****** is true.

Even then, there seems to be no reason for thinking that such a comparatively high confidence in future success should be thought to be something definitive of action. An alcoholic might have more confidence that they will end up having relapsed in a situation involving the presence of alcohol than “otherwise”, on most of the interpretations of “otherwise” we can think of. But even if the alcoholic will relapse, they might not yet be, and hence the kind of attitude described by *Practical Belief****** does not seem to embody the specific kind of understanding an agent has of what they are doing at the time of action, much less a kind of attitude through which acting might somehow be defined.

Reconciling Practical Belief and Failure-tolerance

What seems to emerge from the previous discussion is that, in trying to define action as the object of practical belief, we need to think of the belief in question as thoroughly imperfective. When N is acting in Aing, N believes that they are *Aing*, where this is not to be understood as a piece of knowledge, belief, or comparatively high confidence concerning any past or future success in doing A or anything else. But as also emerged, this does not mean that these attitudes, oriented towards future or past successes, have *no* place in an account of what it is to act. Perhaps some belief about past success, some anaemic hope of future success, and even some comparatively high confidence in such success, are all preconditions for action.

Practical Belief and *Failure-tolerance* may now seem to be neatly and unproblematically compatible. For endorsing them allows us to say, for example, that someone who is acting in providing for the future welfare of their children must believe that they are doing that, although such welfare may not ultimately have been provided. Similarly, it allows us to say that someone's acting in pumping poison into their neighbours' well requires their awareness of such pumping, which can exist even if, due to bad piping, no poison will arrive at the intended destination. So the above discussion gives us a direct answer to the initially worrying question of whether and in what sense a practical belief could be true even if nothing were there to verify it: In a straightforward sense, practical belief *does* require its object, since it is only true if its subject is on the way to Aing, or engaged in Aing, or in the process of Aing, or, for short, *Aing*. But there is something notable in the vicinity that a practical belief does *not* require, since as such it is not a belief to the effect that its subject will end up *having A-ed*.

But there is a subtle complication. In chapter 1.2, I observed that where

a subject is doing an *achievement*, future success in doing it is guaranteed. If someone is, say, walking all the way to school, then they must end up having walked all the way to school. If, now, we take an achievement, and suppose that a subject is *acting in doing* it, we have the makings of a challenge to *Practical Belief*. This seems to be precisely the sort of challenge that is pursued by Davidson:¹⁴

It is a mistake to suppose that if an agent is doing something intentionally, he must know that he is doing it. For suppose a man is writing his will with the intention of providing for the welfare of his children. He may be in doubt about his success and remain so to his death; yet in writing his will he may in fact be providing for the welfare of his children, and if so he is certainly doing it intentionally. (Davidson, 2001a, pp. 91-92)

Davidson also provides this different example:

[I]n writing heavily on this page I may be intending to produce ten legible carbon copies. I do not know, or believe with any confidence, that I am succeeding. But if I am producing ten legible carbon copies, I am certainly doing it intentionally. (Davidson, 2001a, p. 92)

It seems clear that, in both cases, the challenge rests on supposing that the subject is doing an achievement. For the reason why the will writer doesn't know that they are providing for the welfare of their children must be that this requires that (in the future) these children's welfare ends up being provided for. And the reason why the carbon copier doesn't know they are producing ten

¹⁴As before, the difference between knowledge and true belief is not pertinent, and Davidson's talk of doing something intentionally may be translated into talk of acting in doing something.

legible carbon copies must be that this requires them to end up having produced ten such copies.

We seemingly cannot make such challenges by using doings which are not achievements. For what is the problem about saying, for example, that the will writer believes they are pressing very hard on the paper with the intention of producing ten carbon copies? Equally, there seems to be no Davidsonian problem about saying that someone who is acting in walking towards school (and hence doing an activity) believes they are walking towards school, or that someone who is acting in walking to school (and hence doing something of the paradigmatic sort), must truly believe, or know, that they're doing that. The Davidsonian doubt is more like the doubt that someone acting in *walking all the way to school in time for the first lesson* (and hence doing an achievement) must believe they are doing that. And the reason for Davidson's doubt – similar to the doubt attaching to the previous two achievements – seems to be that it is not clear that someone doing such an action must be in a position to accurately predict that they will be successful in carrying it out.

Having noticed that Davidson's challenge seems to stem from his registering how achievements require future success for present engagement, and a concern about saying that an agent is necessarily able to predict future success, we may express his challenge in the form of the following more abstract argument:

Achievements For some of the things N can be acting in doing, believing N is doing it requires believing N will end up having done it

Non-prediction For *all* of the things N can be acting in doing, N can be acting in doing it *without* believing N will end up having done it

Non-belief (C) For some of the things N can be acting in doing, N can be doing it *without* believing N is doing it

This abstract version of Davidson's argument has the troubling conclusion that, after all, it is possible to be acting in doing something without believing one is doing it. Since the argument seems valid, we must find a way to reject one of its premises if we're to stick with *Practical Belief*. But on reflection, both premises are more vulnerable to criticism than Davidson makes apparent.

A very straightforward way of responding to the argument would be by denying its first premise. This denial might be grounded by the simple suggestion that achievements are not, as such, things that someone can be acting in doing. The suggestion would be that no one is ever, as such, *acting in* winning a race, even if, perhaps, it is possible to be winning a race, or at least to have won it (and, though it might seem like a strange claim to take on, it might even be true that someone can, after the fact, *have been* acting in winning a race). If achievements are understood as things which no one can be acting in doing, then even if believing one is doing an achievement (like winning a race) entails believing one will end up having done it (ending up, therefore, having won it), this will not cause any trouble for *Practical Belief*. A different way of grounding a denial of the first premise would be reverting to the more or less Rylean line (back in chapter 1.2 on page 13) on which nothing is ever *doing* an achievement.

If someone should object that this response leads to a too restrictive conception of action, it is worth noting that, wherever someone might be tempted to say that someone is acting in doing an achievement, it seems very possible to say, instead, that they are acting in doing something which, if things go well, will result in them having done an achievement. Hence instead of saying someone is acting in winning a race or in securing the favour of their uncle, we may say that they are acting in participating in the race — perhaps very successfully or ferociously — and may therefore end up winning it, or that they are acting in blandishing their uncle, and may well therefore end up securing the uncle's

favour. And this way of describing these proceedings is really no less natural than that which it would replace. It is arguably more natural than it.

A somewhat less straightforward response to the Davidsonian argument denies *Non-prediction*. The response would have it that, where someone is acting in doing an achievement, they *do* have a predictive, true belief about future success, although in other cases such a predictive belief is not required for acting. This response to the argument is friendly to *Practical Belief* — it merely adds to this that, because sometimes subjects act in doing achievements, something along the lines of *Practical Belief** must also be true of those cases. But the response seems to come at the cost of needing to qualify *Failure-tolerance*. For once some actions are counted as achievements, we will have to make room for the idea that, for some of the things a subject can be acting in doing, present engagement does entail future success. This is not, however, a very great cost, as a properly restricted version of *Failure-tolerance* must still be true: Present engagement in an action does not, *as a general matter*, entail future success in doing it. That is the moral of the case of the builder, and of a large number of ordinary cases of acting, and this is still seemingly enough to undermine *Practical Belief** in its fully general guise.

So once we note that the Davidsonian challenge is limited in scope, and rests entirely on the idea that achievements can be actions, this challenge is blunted. For we may either construe achievements as a special case of action, making special provisions for this kind of forward-looking action, without compromising *Practical Belief*, or else we may simply rule out the claim, on which Davidson's entire challenge rests, that achievements can be actions, and find other ways of talking about an agent's participation in bringing about an achievement. If either response is workable — and on the surface both seem to be — Davidson is unsuccessful in undermining *Practical Belief*.

3.4 Practical Belief is Causal

But having construed practical belief imperfectively, so that it does not require its subject to be successful in whatever it undertakes to do, and having realised the limits to the Davidsonian challenge, there is still, I think, a further challenge to *Practical Belief*, which rests on a further kind of counterexample. The challenge, which I think might be traceable to Anscombe, is expressed by Setiya, who at least *mostly* formulates it in imperfective language, and *apparently* thinks of it as arising outside the case of achievements:¹⁵

Consider a [...] case of recent paralysis in which, at a certain point in my recovery, I am cautiously but not irrationally optimistic: I think that I might be able to clench my fist. Once again, things happen to work out. I clench my fist, and I do so intentionally. Still, given my doubts, I do not believe that I am clenching my fist — perhaps I cannot feel it, or see it, and I am not at all sure of my ability.
(Setiya, 2008b, p. 390)

Let's not construe this as an instance of the simple challenge we started with, as the language in it suggests that the belief in question concerns something which the subject *is* doing, not their future success in doing it. And let's not construe "clenching" as a verb of achievement. That would seem to turn Setiya's worry into an instance of the Davidsonian objection above, but Setiya's challenge seems distinct from it. Rather, let's take it in the straightforward way in which it seems to be intended: Someone is acting in holding their fist closed for a while, or perhaps in moving their fingers towards their palm, where this is the sort of undertaking that can end up having failed or succeeded in the future.

¹⁵Like me, Setiya accepts a version of Anscombe's contention about the role of practical belief in action, and wants to develop it in the face of various counterexamples. But as has already been seen, his way of developing it to avoid the counterexamples is quite different from mine.

Setiya seems to suggest that even when we understand the case in this way, there is reason to doubt that this subject *believes* they are doing the action in question.

But it should be noted that, once practical belief is construed imperfectly, such examples do not in any simple and straightforward way disprove *Practical Belief*. The fist clencher may not think that they are going to succeed, nor that the clenching is going particularly well. Even so it is hard to deny that this subject, if they are acting in clenching, must have some kind of awareness directed towards the fact that they are clenching their fist. Setiya seems right to think that there is some intuitive pull to the idea that a subject like this can act in clenching although, due to some epistemic misfortune, they don't believe they're clenching. But it is not at all clear what it is in the case that is supposed to support this intuition. What is the reason for thinking that, in a case like the one described, someone can act in clenching without believing they are clenching?

I think that the line of thinking which is supposed to bring us to doubt that the paralysis victim believes they are clenching has the following shape:¹⁶ If it is true that, wherever someone is acting in doing anything, they have a practical belief which represents them as doing that, and if we are to take seriously that, for example, Setiya's paralysis victim is acting in clenching their fist, then we must say that this agent has a true practical belief which represents them as clenching their fist. But if we also take seriously that this agent does not believe that their hand is moving in the way that is characteristic of such clenching, then the content of the agent's practical belief (which we are supposing is true) can't be that such movements are happening. But then we may ask, what *is* the object of this agent's practical belief about clenching their fist, if not (at

¹⁶Unless I'm mistaken, this or closely related trouble was discovered by Anscombe (1963, p. 51).

least something that requires) that such hand movements are taking place?¹⁷

It seems to be this type of challenge which is supposed to provide a reason for doubting *Practical Belief*. And of course examples of this sort of challenge could be multiplied. They are in *Intention*: Someone acts in pumping poison into their neighbours' well, but they have doubts about the piping and hence can't vouch for the poison's progress, so in what sense do they practically believe they are pumping poison? Or someone is building a house by giving orders, but isn't sure whether their orders are being implemented, so in what sense do they practically believe they are building a house? If we insist that *Practical Belief* is true, then we seem committed to finding some general way of answering such questions. And if it turns out we can't provide such answers, then that would at least give some indirect support to Setiya's troublesome intuition, putting *Practical Belief* into serious question.

Anscombe warns against two very direct but deeply problematic responses to such questions: The first pronounces that the real object of practical belief is just that, in some very interior sense, one wills something to happen, where belief in this willing doesn't require any belief that hands, tables, tongues, money or anything else moves or does anything else (Anscombe, 1963, pp. 51-52). The second, more obscure, but for many purposes similar suggestion, is that thinking one is acting in doing something is all there is to acting in doing it (Anscombe, 1963, p. 52).

As Anscombe notes, both accounts are troublesome. Both seem to make practical belief infallible (the former by regressively "pushing back" the object of the belief so that apparently nothing could interfere with it, and the latter by

¹⁷It seems that it is "the difficulty of this question" (Anscombe, 1963, p. 51) which ultimately leads Anscombe to the idea that knowledge of what one is acting in doing is different from knowledge of what one is doing in some non-action-involving sense (Anscombe, 1963, pp. 88-89), possibly, but not clearly, with the addition that these two "knowledges" have different objects. (But as will soon emerge, the object of practical knowledge is not a mere willing or something entailed by belief in it for Anscombe.)

analytically legislating such infallibility). Relatedly, both result in a too narrow conception of the types of action a subject can do, since they make subjects unable to act in doing anything we might vaguely describe as “world involving”, like calling an ambulance, taking a walk, or clenching a fist. The quite ponderous worry which underlies both of these, which is difficult to put into clear language but is worth pointing towards, is that both seem to make the object of practical belief void: It is something which can hardly be distinguished from nothing.

Our question hence becomes: What could the object of a practical belief be if, on the one hand, it is not a kind of process which — in unusual circumstances, perhaps involving paralysis or lack of sensory feedback — its subject might not believe is happening, *nor* a special kind of process belief in which is somehow guaranteed, like a purely interior episode of willing, or something mysterious the reality of which follows from belief in it?

For the beginnings of an answer to this question, consider that, where someone acts in doing something, there will tend to be other things they are doing which in some intuitive sense are phases, parts or stages of that comparatively large project.¹⁸ If someone is clenching their fist, this may involve the phases, say, of making an effort to clench, muscle flexing, finger movement, palm encounter, and keeping one’s hand clenched. For present purposes, the most important feature of such phases is that they need not temporally coincide. From a temporal point of view, they can happen in sequence (someone clenches their fist by moving their fingers halfway to their palm, and then the other half of the way) or overlap (someone is baking by putting the oven on while kneading a dough) or fully coincide (someone is lifting a heavy stone by

¹⁸We should not be tempted to think that there is a particular number of phases for every case where someone is acting in doing something. It is no more or less justified than the above to say of this fist clenching that it involves some fingers moving halfway towards the palm, and the rest of the way, and resting in that palm (making for three phases), or just that the subject is making an effort to clench, and clenching (making for only two phases), or that the subject is making an effort for a few moments, and making an effort for some further moments, and starting to clench, and keeping at it (making for four phases).

struggling with their arms and legs).¹⁹

This seems to show that acting in Aing is being in some phase or other, but not necessarily all, of Aing. But of course then *believing* one is acting in Aing should amount to *believing* one is in some phase or other, but not necessarily all, of Aing. This immediately suggests a response to our problem, since it allows us to say that a practical belief which represents its subject as clenching their fist can exist even though this subject does not represent themselves as moving their hand, so long as this subject does represent themselves as flexing their muscles, or holding their hand clenched, or as doing something else which is a phase of their clenching their fist.

Any more radical putative counterexamples, where subjects do not even believe they are in some phase of their action, seem to fall to the ground, since there seems to be no reason for saying that such subjects are acting in doing what they're doing. If we try to imagine that Setiya's paralysis victim is totally in the dark about whether any phase of clenching is happening, it just seems misguided to claim that this subject is somehow acting in clenching, even if this subject hopes that their hand moves and it does.

This solution to the problem obviously rests on some conception of what it is for something that happens to be a phase of a given action, which though somewhat intuitive, has not been elucidated. This conceptual issue might be pressed in a sceptical manner, by presenting a putative counterexample to the thesis and asking why the account should not apply to it. Next I will do this in a way that prompts the introduction of one criterion for phasehood.

Imagine two paralysis victims, L and M, who decide to try to clench their respective fists without being sure whether their respective paralyses have worn off. Imagine that L succeeds and M fails. We've said that if L believes that L is

¹⁹For a more in-depth account which seems to lend support to mine, see Stout (1996, pp. 46-62) on processes.

acting in clenching in the sense that L is in, say, the effort phase of clenching, and L is in the effort phase of clenching, then L's practical belief is satisfied, and L is acting in clenching. So we have given the following rationale for the claim that L has a true practical belief, even in a case where L does not believe their hand is moving:

- L believes L is acting in clenching in that L is in the effort phase of clenching
- L is in the effort phase of clenching
- So L is acting in clenching

Now for all we've said and all we can think to say, it seems possible for M to believe just the same sorts of things that L believes, and hence M may believe that M is making such an effort too. And isn't it possible for M to be making an effort to clench their fist just as much as L, though M unfortunately fails? If we go along with this, we have a problem, since if we go along with it, we can readily construct the same rationale as above for M's case:

- M believes M is acting in clenching that M is in the effort phase of clenching
- M is in the effort phase of clenching
- So M is acting in clenching

And if we accept this rationale, trying utterly unsuccessfully to clench is sufficient for acting in clenching, and then we are seemingly back, or at least on the way to, a version of the sort of view we just declared undesirable. If we are not to accept it, we are committed to finding some as-yet-unelucidated difference between L and M which shows us why M isn't acting in clenching. I said I don't see why M couldn't believe what L believes. But could it be that L but not M is in the effort phase of clenching? A defence of this answer requires some criterion

for phasehood. And I propose the following: For something that happens to be a phase of an action requires that it causally contributes to that action.

On this criterion, the reason why M isn't in the effort phase of clenching is that, whatever M is doing, and whatever beliefs M has, M is not doing anything that causally contributes to their clenching their fist. This criterion gives our account the right implications, in that, on the one hand, it allows people to count as acting in Aing even if they aren't sure that they have entered some particular phase of Aing, whereas, on the other hand, it does not allow people to count as acting in Aing if they don't truly believe they are doing something which causally contributes to their Aing. I do not see how our account could have these implications without containing a causal criterion, since I do not see any non-causal but pertinent difference between cases like L's and cases like M's.²⁰

3.5 Practical Belief is Part-Homogeneous

I wondered: if the kind of practical belief constitutive of acting in Aing represents its subject as acting in Aing, is it possible to go beyond this schematic specification of the object of that belief? So far, I have argued that it must represent its subject as *Aing*, in a sense such that Aing is being in *some phase* of Aing, where being in some phase of Aing turns out to require doing something that *causally contributes* to Aing. In this section I want to argue that believing one is acting in doing something involves believing not just that one is doing further things that broadly enable, promote, or result in it, but also that one is *acting* in doing the things that thus contribute to it.

²⁰See Wolfson (2012, p. 330) for some claims which anticipate some of the present claims in a very rough way. The main differences between Wolfson's claims and mine seems to be that he favours the notion of something's culminating in something else over my notion of something causally contributing to something else, and that he seems to view what is on the far end of the culmination "relation" as a perfective fact about what something has successfully done, rather than viewing it as another imperfective fact.

We may sum up the our developing hypothesis as follows:

Causal Practical Belief: For N to believe that N is acting in Aing requires N to believe that N is Xing in a way that causally contributes to N's Aing.

It may seem that, if there were nothing to the belief of the sort “N is acting in Aing” but a belief of the sort “N is Xing in a way that causally contributes to their Aing”, then we would have in this observation the makings of an exhaustive account of the object of practical belief, which would roughly just replace the above “requires” with “is”. It may also seem that such an account would dissolve our earlier anxieties about the circular shape of our account preventing an informative account of the object of practical belief, since now we would be in a position to say: A practical belief just represents its subject as doing things which cause things, where all this is expressible in action-neutral terms!

But the vague kind of account imagined is really ambiguous between constituting a genuine but false account and a non-account. Hence any reductive hopes of the sort sketched must result from a cross-eyed reading which does not appreciate this ambiguity.

For notice how, so far, we have been using the letter “A” as a stand-in for anything a subject can do. If we read the “X” just invoked as also a stand-in for anything a subject can do, then saying a belief about acting in Aing amounts to a belief about Xing causing Aing is quite obviously false, since someone can clearly believe they are doing something (like sitting handcuffed on and to a bicycle) that causally contributes to other things they are doing (like nearing Kathmandu) without thereby believing they are acting in doing the latter thing.²¹ Metaphorically, this form of account asks the notion of causal

²¹This looks similar to a kind of thesis Anscombe warns against, albeit the warning is presented as a linguistic observation and is addressed to slightly different concerns:

Say I go over to the window and open it. [...] Someone who hears me moving calls out: What are you doing making that noise? I reply ‘Opening the window’. [...] But I don’t say the words like this: ‘Let me see, what is this body bringing

contribution to do all the work in accounting for the content of a belief about acting, and that notion is simply not up to the task.

If, on the other hand, we read the “X” not as a stand-in for anything a subject can do, but as ranging over some more restricted set of things a subject can do, then until that set is specified, the thesis we are considering is not an *account* of the object of practical belief, but at most a promise of an account. This promise is not undermined by observations to the effect that it is possible to have a belief of the general shape “X is sitting handcuffed on this bicycle, which promotes their approaching Kathmandu” without thereby believing X to be acting in approaching Kathmandu. But its promise will be undelivered unless we can form some determinate conception of which kinds of doing contribute to someone’s doings when these are actions.

I think it is possible to deliver on this promise, but only by saying that the doings which figure in the thought of an agent as the causes of their actions are themselves actions. If we do not allow ourselves this, we will have to try to give an account of the following shape:

Event-Causal Practical Belief: For N to believe that N is acting in Aing is for N to believe that N is Xing, where this doesn’t entail acting in Xing, and where Xing is some particular type of happening (...), and doing it causally contributes to N’s Aing.

But it seems impossible to make any headway with this kind of account. The trouble can be gleaned by taking a particular example. Say someone thinks they are acting in pushing a stroller to their home. Of course such a subject might think that they are doing other things, which they do not obviously need to view as their actions, which contribute to their pushing of the stroller, such as being alive, or holding the handle with enough force and friction for it not to

about? Ah yes! The opening of the window’. (Anscombe, 1963, p. 51)

roll down a slope. If the subject is an anatomist, they might know all about the fine details of bodily processes that support and enable their strolling, but even if not, they might know that things inside their body make their body push the stroller with enough force to move it. Even if some such beliefs, which pass some difficult-to-formulate test of not representing its subject as acting, can or must be forced into existence by the presence of a practical belief, it still seems impossible for such beliefs to jointly make up any belief to the effect that its subject is acting in doing anything.

To the same extent that this is convincing, we should endorse the following thesis:

Action-Causal Practical Belief: For N to believe that N is acting in Aing requires N to believe that N is acting in Xing and that this causally contributes to N's Aing.

This account overcomes the difficulties of the previous, and explains those difficulties. For the type of belief that we are willing to say represents its subject as being in a (perhaps very early) phase of acting in clenching their fist, or pumping balloons, or pushing a stroller, is seemingly always a belief which represents its subject as *acting* in doing something else that contributes to clenching, pumping or pushing.

3.6 Practical Belief is Instrumental

Some critics of Anscombian,²² but also Anscombe herself,²³ think that it is possible to think one is acting in doing something that causally contributes to something else which, nevertheless, one does not think of oneself as acting in

²²Recently notably Paul (2011) and Bishop (2011, p. 219).

²³I have in mind the passage involving someone who is pumping poison into their neighbour's well but does not "care tuppence" about poisoning them (Anscombe, 1963, p. 42).

doing. If they are right, we can't flip the above "requires" to "is".²⁴

It would simplify matters if we could respond to this problem by denying its premise, and insist that viewing something as an outcome of one's action is viewing that too as one's action, albeit perhaps as an outcome one doesn't care for or about.²⁵ In fact it is quite easy to produce a counterexample to such insistence: Imagine a subject knowingly acts in playing bowling, and discovers that the ball is about to hit someone's foot because they have suddenly stepped into the lane. They will thereby believe they are acting in playing bowling and that this causally contributes to someone's foot being injured. But of course this doesn't make it true that the bowler thinks they are acting in generating that foot injury. It does not seem right to insist, at this point, that the bowler is acting in injuring, but just doesn't like it, since the reason for our judgement is *not* that the bowler is unhappy with that outcome (they might have thought it to be a funny accident), but something we might vaguely express by saying that the bowler doesn't view the causal relationship between action and outcome to be of the right sort to confer actionhood on the outcome.²⁶

There seems to be a broader and a narrower way of thinking of the causal contribution one action makes to another thing one does, and it seems that only the narrower guarantees that one thinks of the other thing as one's action.

²⁴Following Aquinas, many think that, when some happening is merely foreseen by a subject, this has different ethical implications from when it is that subject's action. If that were true in general, we might have been able to resolve our action-theoretic question by thinking about differences of ethical implication between some well-selected cases. Perhaps some are engaged in this very project. But absent a defence of the conditional — and I can think of none — this strategy seems flawed.

²⁵As Paul (2011, p. 13) points out, Castañeda can seem to have said something in defence of such view, since he has said that foreseen byproducts are "endorsed" just as much as intended consequences (1979, p. 255). But it is not clear-cut, since he has added that though both are endorsed, foreseeing and intending are distinct "determinates under the determinable endorsingly thinking" (Castañeda, 1992, p. 452).

²⁶Notice that we could not have produced such a counterexample if we were still in the grip of a view of practical belief as a perfectly oriented, predictive belief. For on such a view we could have replied that the subject did not make any predictions about foot injuries right at the start, when the action was set in motion. Since by now we understand that practical belief must be such as to accompany and be verified by its object at the time of acting, we need to respond to the problem by thinking more carefully about what its object is.

If a subject thinks merely that they are acting in Xing and that this causally contributes to their Aing, they may not think of themselves as acting in Aing. But if a subject thinks that they are doing X *in order to* A, they seemingly must. This tempts us to define the object or practical belief as follows:

Instrumental Practical Belief: For N to believe that N is acting in Aing is for N to believe that N is Xing *in order to* A.

This is, I think, a healthy temptation. But it can look as if we now have to provide a definition of the connection signified by “in order to” as it figures in the subject’s practical belief. Sarah Paul has recently charged accounts which attempt to understand the psychological item that defines action “solely as a kind of cognitive, belief-like grasp of what one is doing” (2011, p. 12) with a problem of explaining the difference between the (cognitively grasped) things one does which are actions from those (seemingly equally cognitively grasped) things one does which are mere foreseen side-effects. She very briefly considers the move, which we are presently performing, of explaining the difference by saying that the contents of the beliefs in question are different, and that foreseen side effects are not considered as things one does other things in order to do. She complains of this move as follows:

But it is essential to see that it is not an answer to the problem simply to help oneself to the notions of “aim,” “means” or “end,” and read the structure of what we intuitively take the agent to have done back into the agent’s thought. We are after a theory that explains *what it is* to treat a state of affairs as an aim as opposed to a byproduct. (Paul, 2011, p. 13)

If I understand Paul, the issue she raises is that though we may have an intuitive idea of a specifically instrumental causal connection between X and A, and

while it may be definitive of acting in Aing that a subject thinks they are doing something in order to A (as opposed to just doing something which causes Aing), this does not provide for a satisfying account until we can explain the nature of such specifically instrumental connections as they figure in an agent's thought.²⁷ (Note the very clear affinity between this complaint and the one quoted from Peacocke on page 71.)

How can we define the kind of causal connection a subject imagines holds between two happenings when they are doing one in order to do the other? On comparing cases where things cause things to happen with cases where agents do things in order to do other things, we find no pertinent difference in the causal routes, or the items involved in them. Hence it seems to be a mistake to introduce thoughts about such differences to try to account for the difference between a thought about causes and one about instrumental causes. But in fact I think our answer is in our question: When a subject does one thing in order to do another, they act in causing one by doing the other.

Thus, if our problem was finding some way of eliminating the “in order to” below —

Instrumental Practical Belief: For N to believe that N is acting in Aing is
for N to believe that N is Xing *in order to* A

— then it would seemingly be solved by the following:

Instrumental Practical Belief*: For N to believe that N is Xing in order to
A is for N to believe that N is acting in *Xing so that it causally contributes*

²⁷It might be thought that there is a ready response to this, in that when someone does X in order to A, what this comes to is that X and A constitute a “unity” which “makes an intentional action out of them” (Thompson, 2008, p. 132). Suggestive as this sounds, and ignoring the question of what unities-in-general are, this will have us asking which kind of unity makes an intentional action out of its elements. Thompson seems to acknowledge that specifying the requisite kind of unity will require a philosophical account of “the intellectual aspect” or thought that is involved just when someone does something in order to do something else (Thompson, 2008, p. 133). This is what I want to provide next, though admittedly there is something bare or schematic about my specification of this thought.

to Aing

Instrumental Practical Belief responds to Paul's original problem of distinguishing actions from foreseen side effects: Foreseen side effects aren't covered by the agent's practical belief, since this belief represents its subject as doing something in order to do what's done, and since nothing is done in order to produce a side effect. *Instrumental Practical Belief** then responds to her further worry that we have no hold of the content of that attitude: For a subject to think of themselves as doing something *in order to* do what's done just is for this subject to think of themselves as *acting* in doing something that causally contributes to that.²⁸

Is it unhelpful to respond to Paul's two worries by introducing two theses that are flagrantly interdependent in this way? Since Paul's original complaint seemed to be a complaint about introducing analysans in analysandum, and since it seems that we are still doing this, it can seem that Paul would not be satisfied with the response. Since Paul is not explicit about the grounds of her complaint, it is not obvious just why she would object to it. I can think of two sorts of grounds for complaint: It might be that Paul is simply demanding a *reductive* account of the notion of doing something in order to do something else, which figures in a practical belief. But since this notion seems indispensable for action theory, and since it does not seem that such an account could be given, this does not seem to be a good demand to make. Alternatively, it might be that Paul's demand is not a general demand for a reductive analysis, but a version of the more ponderous but less clear kind of worry which we set aside right at the

²⁸"Xing so that it causally contributes to Aing" thus signifies a complex proceeding in which N is Xing, and where N's Xing is causally contributing (however indirectly) to Aing, as when, for example, someone's planting a seed is causally contributing to their improving next year's crops. This notion of a causal contribution, when taken on its own, is not simply a cloaked way of referring to an instrumental connection. The suggestion is rather that thinking of someone as *acting in effecting* such a causal contribution between X and A, *is* thinking of them as Xing in order to A, and vice versa. (This should not be taken as suggesting that either notion is somehow more "basic" than the other, or that one might be reduced to the other.)

start: If Aing in order to B (so that Bing is not a mere unintended side effect) is accounted for by saying it involves believing that one is Aing in order to B, then how are we to give an account of the content of that belief? Will this not now be impossible? I will return to that form of objection in the fifth chapter, but before then I want to examine the consequences of the account that has now fallen into place.²⁹

²⁹The mere fact that only happenings that a subject thinks of themselves as acting in doing can figure as happenings they think of themselves as doing in order to others should not generate metaphysical trouble (“what is this queer and unscientific sort of causal connection that is proper only to action?”) or epistemic trouble (“how does a subject know that they are *acting* in making pancakes, since all they encounter in the world is that arms, flour, milk, eyes, and a spatula are moving around in such-and-such a way!”). For grasping any relation, including that signified by “... weights as much as ...”, “... is richer than ...”, or “... is made of cucumbers and ...” requires a grasp of the kinds of item which can be related through it, and this does not tend to make these relations, nor knowledge of them, seem mysterious. If there is trouble, it is, I think, generated by the additional fact that grasp of the items that fit in the “in order to”-relation requires grasp of that relation, as it probably doesn’t in the case of the “... is made of cucumbers and ...”-relation.

Chapter 4

Consequences

4.1 Introduction

The present chapter is devoted to examining three consequences of the account on offer. The first consequence is that it is impossible to act without taking some other action as a means, and thinking of that action as one's means. Hence the notion of a basic action, as well as the notion of a basic practical belief, is rendered impossible. In defence of this consequence I undermine a causal regress argument meant to secure the need for basic action, and also some claims about the limits of an agent's thinking which are meant to secure the need for basic practical belief.

The second consequence of the account is that, when coupled with natural claims about action explanation, due to Ryle and Davidson, it renders every action explainable by appeal to an agent's desire to do an action of that kind. This defends, as has proven difficult to do, the Anscombian contention that the topic of action is constitutively subject to a special mode of explanation. But it meets with an objection, found (intertwined with others) in Nagel, to the

effect that desire ascriptions could not be explanatory if their applicability is guaranteed. But I argue that this objection is mistaken, by appealing to the fact that there is always more complexity to someone's desire than is revealed by the mere fact that they are acting, so that such complexity might always be appealed to in a genuine explanation of an action.

The third consequence of the account is that it affords a way to combine two thoughts which have recently been assumed to be in conflict: One thought is that the objects of our desires are normally revealed right on the surface of our ascriptions, but the other is that desire ascriptions are often in some sense incomplete, revealing only part of what the agent really wants to do. I reconcile this pair of thoughts by explaining how a desire ascription sometimes specifies a part of an instrumental chain of desires.

4.2 Basic Action

One way of expressing the results of the previous chapter is as follows:

Simple Practical Belief For N to be acting in Aing is for N to truly believe that N is acting in Aing, where

Instrumental Acting in Aing = Xing in order to A

Causal Xing in order to A = Acting in (Xing so that it causally contributes to Aing)

From this way of putting things, it is obvious that the present thesis generates regresses. One kind of regress runs from action to action: Acting in doing anything requires acting in doing something else which causally contributes to that, so acting in doing anything requires acting in doing an infinitely unfolding set of further things, and, moreover, causal connections between all the actions

thus performed. Another kind of regress runs from instrumental belief to instrumental belief: Since believing one is acting in doing anything requires truly believing¹ one is acting in doing something else which causally contributes to that, believing one is acting in doing anything requires believing oneself to be doing an infinitely unfolding set of further actions, and, moreover, believing in causal connections between all the actions thus believed to be performed.

Our question now becomes whether these regresses are vicious. Some, including Danto (1965), have argued that the first sort of regress proves vicious, and others, including Hornsby (1980), have argued that of the second kind. These charges have motivated, respectively, the introduction of *basic action* — a kind of action which nothing is done in order to do —, and what we might call “basic practical belief” — a belief that one is acting in doing something which requires no conception of a means by which one is doing it.² Either charge of vicious regress seriously threatens the results of the previous chapter, since if one can act in doing something, or believe one is acting in doing something, without acting in doing something else, or believing one is, then instrumental connections of the sort just sketched would not be definitive of action. (They might still be definitive of an important *kind* of action, but reverting to that line would raise a troublesome question of what unifies this kind with others.) But in fact the charges of vicious regress rest on weak arguments, as I will next argue.³

¹If acting in doing something only required a belief (true or false) about doing some further action in order to do it, then it could be argued that infinite beliefs would not be generated. But the above thesis requires, for doing any action, that there is a *true* belief which represents the doing of a further action. So it obviously does generate infinite beliefs.

²Basic action actions are reasonably thought of as those requiring only basic practical beliefs, and basic practical beliefs as those proper to basic actions, as it would be deeply uncomfortable (by anyone’s standards, I think) to promote only basic practical beliefs (“N is doing X in order to A, they just have no idea — to them, it’s just a matter of acting in Aing!”) or only basic actions (“Silly N *believes* N is doing X in order to A, they just happen to be wrong about that — in fact they are doing nothing in order to A! They’re *simply acting in Aing!*”)

³Though I think the issue would be worth raising, my discussion below focuses less on the intelligibility of the basic entities introduced to stop the regresses than on the regresses that

The first kind of vicious regress argument, and the resulting introduction of a basic kind of action, was made prominent by Danto. But curiously Danto's original formulation of the regress does very little to explain why it is vicious. Considering a case where someone named M is acting in pushing a stone, Danto says this:

[I]n order to cause the motion of the stone, something else must be done, or must happen, which is an event distinct from the motion of the stone, and which stands to it as cause to effect. Now this other event may or may not be a basic action of *M*'s. But if it is not, and if it remains nevertheless true that moving the stone *is* an action of his, then there must be something else that *M* does, which causes something to happen which in turn causes the motion of the stone. And *this* may be a basic action or it may not. But now this goes on forever unless, at some point, a basic action is performed by *M*.
(Danto, 1965, p. 145)

Danto's regress argument for basic action proceeds by asking, for each action performed, whether it is caused by another "distinct" action or not, noting that the series of actions must either be infinite or not, and reasoning from the apparent absurdity of infinite actions to the necessity of introducing uncaused action.

Since we understand instrumental connections so as to require causal ones, we must construe Danto's argument for the existence of uncaused action as an argument for the existence of instrumentally basic action. For if (as we're assuming) Aing in order to B requires acting in Aing in such a way that it causally contributes to Bing, and if (as Danto argues) some actions are not caused by other actions, then some actions must be such that nothing is done are supposed to motivate their introduction.

in order to do them. And hence we can find no comfort in the following sort of response to Danto's argument:

Lavin, who is officially an opponent of basicness, says that, even granting that there is basic action in the sense of an action which poses "a limit to a series of causal dependencies", this does not establish that there is basic action in the more action-theoretically crucial sense of an action which nothing is done "in order to" do (Lavin, 2013, p. 282). As should be obvious, this is exactly the sort of move that is unavailable to someone who, like me, thinks that instrumental connections between actions require causal connections. But Lavin's response seems quite incomplete even when taken on its own terms, since it leaves us in the dark about the conception of instrumental connections on which he thinks they *can* exist without causal ones.

Lavin's only real characterisation of an instrumental connection between actions is this negative one: It does not require a causal connection. Lavin can seem to expand on this point when he says that N can do something in order to A even if N's Aing "cannot be analyzed" (Lavin, 2013, p. 282) as a causal consequence of something N did. But whatever this might be thought to add to the basic negative point only makes that point more confusing: What does it mean for N's action to be analysable as an effect of something N did? It does not seem plausible to say of *any* action that it is an action *just in virtue of* being an effect of another action. For if we consider an episode simply as an outcome of a subject's action, we will never seem to have enough information to call that episode this subject's action.⁴

But perhaps Lavin does not mean to make any such suggestion about analysability. Perhaps he simply means to endorse what I called the basic negative point: There might be causally basic actions — actions not caused by other actions — even if there are no "teleologically basic" actions — actions not done

⁴Remember chapter 1.3.

by doing other actions. Why should we think that Lavin is right on this basic negative point? Let's bring in one of Lavin's supposed illustrations of the thesis. Lavin says that a captain may, for example, be *breaking a seal* in order to push a button, although the captain is doing nothing that causally contributes to their pushing the button (Lavin, 2013, p. 282). But this is hardly a pre-theoretically apparent example of what Lavin wants to prove. Why *not* say that the instrumental connection between breaking and pushing introduces causation? After all, the captain's breaking of the seal enables and facilitates encounter between finger and button, as even Lavin (2013, p. 298, n. 19) apparently grants.

Hence it seems that the contention must stem either from some unargued restrictiveness concerning causation (which may not be the conception which features in Danto's argument, so that Lavin's response is not really responsive to that argument)⁵ or from a withheld conception of the nature of instrumental connections (which, if my argument in chapter 3.4 is successful, must be mistaken). In either case there is nothing in Lavin's response which gives positive sense to the thesis that, even if some actions are uncaused causers, all may be instrumental ends. So Lavin's defence against Danto is a retreat too far, guarding against the argument for basic action by pulling causality out of instrumental connections, thereby leaving instrumental connections in obscurity.

Back to the question of why action should be impossible if it always required a further and instrumentally subservient action. There is, of course, an infinite number of potential suggestions, but it is not so easy to come up with reasonable ones. One of the most natural suggestions is to say that, since every action taken as means to another action needs to be taken (however slightly) before that further action starts to be pursued, then taking infinite means-actions would

⁵For example, it might be Lavin's supposition that pushing five buttons cannot be a cause of pushing 50 if it is, in an intuitive sense, part of the pushing of the 50. I see no problem with adopting a notion of causation which introduces such a restriction, but it is not obvious that Danto's argument rests on any such restricted notion of causation.

take infinite time, which surely would make acting impossible, taking, as acting seems to, finite time.

This might be the kind of suggestion Danto has in mind when using the word “distinct” in the quote above. But here there seems to be nothing wrong with denying the premise. Why think that means always must be pursued *before* their ends? If someone breaks an egg in order to make an omelette, the breaking does not start before the making. Since at least some means do not extend further back in time than their ends, infinite instrumental connections need not extend an action infinitely.

Once we focus our attention on the instrumental connections that reside *within* the time frame of an action we invite a different sort of objection, raised and dismissed by Thompson (2008, p. 109). The objection is that since every action taken as a means needs to meet some minimal spatiotemporal threshold (for reasons having to do with our inability to think about things moving along very small distances or time intervals), and since, if there is an action taken as a means to every action, some actions must be pushed below that threshold (for reasons that are not yet entirely clear), then acting must be impossible, since it would require agents to think about things the spatiotemporal boundaries of which are out of their cognitive reach.

Some such inference may be valid, but there seems to be ample opportunity for questioning the first, and maybe also the second premise. As Thompson noncommittally notes, we can deny the first premise of such an argument by adopting the “high road of insisting that the intuitive apprehension of trajectories involved in continuous intentional action always involves an intuitive apprehension of all of the part[-actions taken as means to] all of them, no matter how small” (Thompson, 2008, p. 110).

And I do not see anything wrong with the high road. If someone knows

they're moving their arm along a stretch — say from the side of their body to a window — they'll plausibly know that this stretch can be divided into arbitrarily small further stretches. If now we take some arbitrarily tiny stretch along which their arm is currently moving, then why couldn't we say that they know they are moving their arm along this tiny stretch too? There seems to be no clear reason for supposing that such a movement, however minuscule, is out of reach to the thoughts of such a subject.

There also seems to be room for questioning the second premise of the present sort of argument. To see this, let's not take the high road for the sake of argument. Let's assume that there are minimum spatiotemporal thresholds for what can figure in an agent's instrumental belief. Why should we suppose that, if there are infinite means actions taken to every action, some of them must be pushed below this threshold?

To try to illustrate how someone might arrive at such a thought, consider a case where someone is acting in pushing a boulder from point A to point Z. Suppose we keep iterating the thought that, if they're pushing it from A to Z, they must be doing it by pushing it from A to F, which they must be doing by pushing it from A to C, and so on. Since in so doing we keep making the means actions shorter, smaller and harder to circumscribe from the macroscopic point of view, it may seem that we're fast approaching the minimum threshold — some stretch of pushing, say from A to B, such that the agent can no longer think of themselves as pushing the boulder along any substretch as a means to doing that.

Even if we go along with this, substretches are not the only place in which we may look for instrumentally subservient actions, and it seems possible to find other sorts of action which are performed as means to pushing the boulder from A to B. For example, the agent might be whispering “you can do it!” to

themselves, hoping this will help boost the pushing, or they might be breathing athletically, or, if such things are allowed (and why not?) they might be acting in staying on a rigid exercise routine, running back for years, for the sake of being able to push such boulders when the need arises, or even just (still — *why not?*) just in staying fit enough to make their current project so much as conceivable. Of course the agent may not be taking any of these particular means in the case at hand, but why suppose that they must not take some such preparatory or promoting means, thereby extending the instrumental chain beyond the boundaries of their pushing of the boulder?⁶

The second type of objection, targeting our claims about the omnipresence of instrumental belief in action, has been made, perhaps most prominently, by Hornsby (1980):⁷

Among the things a person knows how to do, some of them he must know how to do ‘just like that’, on pain of needing to ascribe to him indefinitely many distinct pieces of knowledge to account for his ability[.] (Hornsby, 1980, p. 88)

The idea of Hornsby’s objection, I think, is that if we say that doing an action requires knowledge of how to do it, and if we say that knowing how to do an action requires knowing how to do a further action, which one takes as a means to doing it, then we will have the absurd consequence that an acting subject possesses “indefinitely many distinct pieces of knowledge”. Hence, according to

⁶Thompson does not discuss this kind of response. This is probably because he is less interested in defending the general idea that there are infinite instrumental chains of action wherever there is one action, than the more specific idea that, moreover, these infinite actions are ‘organ-like parts’ of the action to which they contribute (Thompson, 2008, p. 107). An organ-like part of an action seems to be a means taken to an action, but not performed before that action, as egg breaking is a means to omelette making, which does not start before it. If Thompson is only interested in finding means that are organ-like parts, *within* the time frame of every action, the present sort of response does not help him. But since I do not operate with such a restriction, it helps me.

⁷The objection ties into a recent discussion about the relationship between know-how and knowledge-that (Stanley and Williamson, 2001; Hornsby and Stanley, 2005).

this objection, we must say that some of the things a person knows how to do do not require such knowledge of a further type of action — they are things the agent “just knows” how to do.

It is tempting to ask of Hornsby, and others who say similar things, just what is supposed to be the content of this knowledge, which so far has only been characterised negatively. But that is not our present topic, as we are interested in the charge of vicious regress itself. The question we should ask, I think, is why it couldn't be the case that an acting subject possesses an indefinite or infinite number of pieces of knowledge — or as I'll mostly keep saying, an infinite number of true beliefs. (The difference between true belief and knowledge still does not matter for our purposes.)

There could be many different sorts of answer to this question. One answer would be that something in the nature of believing (or knowing) makes it impossible to have beliefs (or pieces of knowledge) without limit. But here I think our view should be quite the opposite:

For example, knowing how to do an appropriately delimited subset of basic mathematics precisely *does* seem to require having an infinite number of basic arithmetical beliefs, such as, for example, that $6 + 4$ is the same as 10, that both of these are the same as $9 + 2 - 1$, that these, even together, are less than 139, and so on. And even if, for some reason, which I do not think I would accept, it is judged that the number of beliefs needed to know this basic chunk of mathematics is not infinite, it is very unclear why there should be any deep problem about saying so.

Similar things could reasonably be proposed about basic spatiotemporal understanding: Perhaps if someone knows enough about how size works, there's no end to the number of comparative size judgements they'll be able to make. And perhaps the same thing is true about good taste, so that if someone has

enough of it, there will be no end to the possible combinations of flavour or notes on which they'll be able to give reasonable verdicts.

A different type of objection targets the psychological realism, or maybe even the psychological possibility, of supposing that there are infinitely many instrumental beliefs present where an agent acts. One objection of this type is that if someone is acting in tying their shoelaces, we cannot keep iterating the thought that they are taking some other action as a means to tying their shoelaces, since beyond some point we will not be able to find a course of action that seems to "pass through the mind" of the acting subject.

This is an example of a type of objection that is fairly common in philosophy. But that sort of objection faces a standard response. For if the concern is that infinite courses of action need not necessarily pass through the mind of someone who is acting in tying their shoelaces, a simple response is that what passes through someone's mind should not be identified with what they believe.

Stanley and Williamson (2001) have recently made such defensive points, citing the less recent example of Ginet (1975). But the defence seems to instantiate Wittgenstein's (1953) and Anscombe's (1963) persistent warnings against trying to understand several classes of psychological fact in terms of things being present to consciousness as objects are to perception — against supposing that, if someone is aware of an intention, knows a rule or language, etc., then some thought or impression on the matter must somehow occur in their mind. Trying to model these psychological features on perception or conscious entertaining is mistaken, and the case of belief is a prime example of the mistake. But if, now, a subject's beliefs about what they're doing need not end where their occurrent thoughts or impressions on the matter do, absence of these occurrent phenomena does not prove absence of belief.

There's a different type of objection, which does not rest on the limits to

what someone can occurrently think about or consciously entertain, but rather on the limits to what is in some intuitive sense “available” to their thinking. Even granting that beliefs must not pass through the mind, someone might ask how it could be that someone who is acting in tying has a true belief *how* they do it, if they are not even in a position to describe how, with prompting and reflection. How can an agent count as believing something which is not, in this vague but intuitive sense, available to them?

Here there is a possible analogy with the case of perceptual experience. For in that case, it can be tempting to move from our difficulties with articulating the complexities of how we tell that something has, for example, the taste of strawberry, to vague conclusions like the following: The taste of strawberry is simply given, just impinges. Or: One knows “just like that” what that taste is like, though one has no idea how one manages to know it.

The resulting claim is mysterious to me, just as it is mysterious to me how a subject could know “just like that” how they are doing their actions. But even someone who does not think there is a mystery here should be suspicious of the inference from current inability to articulate “how we tell”, to conclusions to the effect that there is *no way in which we tell*. Churchland makes vivid that there seemingly can be features by which someone tells that something has, say, the taste of strawberry, even if they are not in a position to say what these features are:

I may indeed be *unable* to specify any sub-dimensions whose peculiar concatenation constitutes the sound of my brother’s voice, or the poppy’s visual orange, or the taste of thyme, or the smell of yellow cedar. But neither could the still-learning child specify, at least at the outset, the taste of sweetness, the taste of creaminess, and the taste of strawberry-ness as constituting sub-dimensions of her taste

of the ice-cream cone, even though those elements were undoubtedly there, and even though she subsequently came to appreciate them.
(Churchland, 2014, p. 47)

Bracketing the question of just what Churchland means by a “sub-dimension”, his idea seems to be this: Yes, it is true that if someone experiences a voice as being their brother’s, they may not be able to say — even with some prompting and reflection — on the basis of what they find that it is their brother’s voice. But no, this does not undermine the idea that, when they do experience a voice as being their brother’s, they discriminate it as such on the basis of some features of this voice.

Discriminating something as such-and-such’s voice on the basis of some of its features apparently requires that these features (be it some characteristic patterns in how the voice’s pitch is modulated, or something about its volume or pace) are *available* to the discriminator. So Churchland’s point requires us to say that, although there is perhaps an intuitive sense in which these features are not available to the discriminator (in that they cannot easily say what the features are, even with prompting and reflection), there is another sense in which these features are available.

In fact this idea is not untoward, since it is not clear how we could possibly make sense of a subject’s ability to tell whether a voice is their brother’s without attributing to them an ability to register those features of a voice which are likely to reveal this. Hence though we have no very detailed account of the two forms of availability we need to make sense of Churchland’s thought, what he says seems to be enough to show that there is nothing conclusive about an argument like “N is experiencing something as having the taste of strawberry ice-cream, but can’t say on the basis of which features of the ice-cream this impression is formed, so there can be no features on the basis of which this impression is

formed”.

Once we see that there is a sense in which something can be available to thought which does not require its subject to be able to describe it — even with some prompting and reflection — we’ll see that, analogously, there is nothing conclusive about an argument of the form: “N is acting in moving their legs (or doing something with their muscles, moving their body, making an effort to move their body, or another of the innumerable possible candidates for an action done ‘just like that’), but can’t say what they are taking as their means to so doing, so there can be nothing which they believe is their means to doing it”.

Just as in the case of experience, someone’s inability to say how they do an action should not be thought to undermine the thought that, when they do this action, some conception of a means by which they do it is available to them. It should not be a convincing case for basic action that there are cases where we have trouble in conveying how we do our actions.

4.3 Action Explanation

4.3.1 Constitutive Explanations

Facing the question of what “distinguishes actions which are intentional” Anscombe famously said that it is the applicability of “a certain sense of the question ‘Why?’” (Anscombe, 1963, p. 9). She seems to have meant that part of what it is for a subject to be acting is for what they’re doing to belong to a kind of occurrence which has a certain kind or form of explanation — the kind of explanation that would answer that kind of why-question.⁸

⁸Ford (2015) develops some such understanding of Anscombe’s project, in part via a comparison with Frege’s work on arithmetical explanation. The material in this section is somewhat indebted to Ford, and is at least *prima facie* compatible with the existence of the connections Ford draws between action and arithmetic, though I have little to say on Ford’s

If this is what Anscombe meant, we may note that there is an important but neglected sense in which action theorists nearly universally practice it. For almost everyone who tries to supply an account of action more or less overtly introduces an *explanatory connection* between an agent's action and some sort of explanans deemed essential to the occurrence of the action. The explanatory relationships introduced may or may not be called causal, are sometimes counted "teleological", and about equally often "normative". The elements deemed proper to function as explanans vary: Sometimes they are said to be psychological items (such as an intention, a trying, a volition, a belief about the good, or a Rylean inclination) and sometimes they are deemed to be things external to the mind of the acting subject (a "normative reason", on various understandings of that term). For all these differences, the overarching tendency is clear: Introduce some element in the dual role of being a necessary part (or at least necessarily mostly a part) of the fact that someone is acting in doing something, and at the same time *explanatory* of their acting as they do.

But having noticed this neglected point of convergence, someone may object to it: Isn't how we explain things a different matter from what the explained things are, and isn't it true that there are many and interest-relative ways of explaining why someone does an action, whereas, if acting has a nature, it ought not to be relative to these contingent interests, so that it would be a mistake to try to define action through a kind of explanation? It might be some such worry that Smith is responding to when he spontaneously declares the following in support of what I think of as the Anscombian assumption:

A belief-desire explanation of a bodily movement is [...], as we might put it, a constitutive explanation of an action[.] Other explanations of actions may be available, but they are all non-constitutive: their

contention that the supposed *action explanation* "for no particular reason" (when asked why one is doodling) corresponds to the *arithmetical explanation* "none" (when asked how many stones one has in one's pocket).

availability is not what makes our bodily movements into actions.

(Smith, 2009, p. 58)

For this idea, that (a certain kind of) explainability by appeal to a belief-desire pair is what *makes* it the case that someone is acting, would of course play a part in justifying Smith's desire to put such explainability at the heart of his account of what it is to act. Stoutland clarifies this contention on part of theorists like Smith by saying that on the "causal theory [...] an intentional act just is an act that has a certain kind of explanation" (Stoutland, 1980, p. 351). If *that* contention is true, then there could be no way of avoiding introducing a certain kind of explanatory connection in an account of what it is to act.

Ignoring the details of Smith's own view, but employing the insight behind his use of the term "constitutive", we may express the strongest version of this apparent guiding assumption of action theory as follows:

Constitutive Action Explanation

- a Whenever a subject is acting in doing something, an explanation of a particular form, with a particular kind of explanans, is true of what they're doing
- b Whenever an explanation of a particular form, with a particular kind of explanans, is true of what a subject is doing, then that subject is acting in doing it
- c The two forms of explanation and the two sorts of explanans mentioned in *a* and *b* are the same

Smith expresses, and Stoutland describes, accounts which try to make good on *Constitutive Action Explanation*, but the commitment is shared by many others working in action theory, if sometimes with ambivalence.

To see this, note first that the kind of theorist who thinks that the possibility of giving a rationalising explanation — one which exhibits a further end or reason for doing something — proves that the agent acted for the sake of that reason or with that end in view, goes along *at least* with *b* above. It is true that such a theorist may take a more hesitant attitude towards *a* and (consequently) *c* above, since it can seem possible to act in a way which cannot be explained by appeal to a reason or end (as perhaps in cases involving doodling or acting on minor odd compulsions). But even so, it lends some support to *Constitutive Action Explanation* that, because this kind of theorist has not described a kind of explanation that is omnipresent in action, they do not really seem to possess a full account of what it is to act. Those actions which do not seem to be performed for the sake of further ends or reasons are precisely those which the present kind of theorist will find it difficult to incorporate in their account of what it is to act. For what could plausibly make these marginal odd compulsions into actions, if they do not share in the kind of explainability which defines the primary specimens?

Then note that the kind of theorist who is liable to view action as the sort of episode that is caused by the presence of some bit of psychology must endorse at least *a* above, on the hard-to-deny assumption that “... is caused by ...” entails “... is potentially explainable through mention of ...”. It is true that such a theorist may then reject *b*, and consequently *c*, on the basis of the previously considered deviant causal chain cases, where subjects seem to be subject to the causal form of explanation invoked without acting. But again, this can seem to lend support to *Constitutive Action Explanation*, since, as was noted in the second chapter, a theorist who invokes such a bland form of causal explainability will face what looks like insurmountable trouble in completing their account by adding further conditions sufficient for action. Again, failure to

conform to *Constitutive Action Explanation* seems to lead to a failure to provide a satisfactory account of action, suggesting that, to do better, an account must do justice to *Constitutive Action Explanation*.

Below I want to show, in outline, how to formulate a conception of a kind of explanation which is constitutive of action in the sense that *Constitutive Action Explanation* describes, and which consequently occupies a Goldilocks niche between those conceptions of action explanation which overload such explanations with normative and ends-oriented thoughts, so that such explanations are not always possible where subjects act, and those that view the explainability at issue as a matter of basic causal relatedness, so that such explanations are sometimes possible where no one acts. Not surprisingly, the form of explanation I favour is one on which an agent's actions are declared to be exercises of the dispositions which are this agent's desires. It thereby views the explanation of action as a species of dispositional explanation. After defending this form of constitutivism about action explanation, I will respond to an objection, due to Nagel, which has it that, if it is guaranteed that an acting subject wants to do what they're doing, mention of such a desire can't explain their action.

4.3.2 Dispositional Explanations

Why do I want to tell him about an intention too, as well as telling him what I did? [B]ecause I want to tell him something about *myself*, which goes beyond what happened at that time. (Wittgenstein, 1953, §659)

Paraphrasing this quote, it asks what the explanatory point is of saying that some action stems from some intention, and answers that it is to provide information about the agent. This simple answer seems unavoidable, but it raises the further simple question of what sort of information it is that such an explana-

tion provides about the agent. If now we identify Wittgenstein's intentions with Ryle's motives, we can see Ryle as providing a further answer to this further question, albeit perhaps Ryle's answer is still somewhat incomplete:

To say that he did something from that motive is to say that this action, done in its particular circumstances, was just the sort of thing that that was an inclination to do. It is to say 'he would do that'. (Ryle, 1949, pp. 92-93)

Elaborating a little, Ryle says this:

The imputation of a motive for a particular action is [...] analogous to the explanation of reactions and actions by reflexes and habits, or to the explanation of the fracture of the glass by reference to its brittleness. (Ryle, 1949, p. 90)

Identifying Wittgenstein's intentions, and now Ryle's motives, with my desires, we may easily see the above as suggesting that, in explaining an agent's action by appeal to their desire, we are making an appeal to the presence of a disposition on the part of the agent, and a disposition to do "just the sort of thing" which the agent is doing. Hence saying that some agent does something out of some desire (or motive or intention) is saying that this agent is the kind of agent that is disposed to do that type of thing, which amounts, in however a limited sense, to explaining their action. Of course the Rylean proposal, as so far presented, is vague in several respects. It leaves out the point that if there's a disposition, there must be some associated condition such that, when it obtains, that disposition is exercised. But this point seems present in the following quote:

The statement 'he boasted from vanity' ought[...] to be construed as saying 'he boasted on meeting the stranger and his doing so satisfies the law-like proposition that whenever he finds a chance of securing

the admiration and envy of others, he does whatever he thinks will produce this admiration and envy'. (Ryle, 1949, p. 89)

Here Ryle seems to say that, where there's explanation by motive (or desire or intention), there is an appeal to a disposition of a sort which is exercised *in some type of condition*. Ryle also seems to say that the relevant condition — the one in which a vain motive manifests — is the presence of some belief about what might count as such a manifestation. And on both these points I think Ryle is right.

However, the above quote also seems to introduce further claims, which are more troublesome. Ryle's apparent idea that one can understand "is disposed to A in condition C" as "is such as to A *whenever* C obtains", is problematic, and especially so if the "whenever" is attributed something of the force of a law of nature.⁹ Ryle, further, is not very whole-hearted about circumscribing motives-as-such so as to form a category of explanans constitutively able to explain an agent's actions.¹⁰ This is natural given that Ryle's primary interests seem to lie in the area of action explanation, not action theory, and since he seems not to draw a very tight connection between the two. But even in giving specifications of individual types of motive, like vanity, in order to exhibit the general point that motive explanations (like "he boasted from vanity" are dispositional, Ryle only gives quite rough specifications which can seem merely provisional. (Must someone who boasts from vanity be the sort of person Ryle describes? And could not someone — someone who tends to talk about themselves when nervous — be similarly disposed, and exercise the disposition, and even boast in so doing, and yet not boast from vanity?) But these potentially problematic features,

⁹See chapter 2.3.2 on page 61 for the obvious objection. But see also Ryle's brief statements about laws contra dispositions (Ryle, 1949, p. 43) and his brief attempts to deal with exceptions to ability ascriptions (Ryle, 1949, pp. 123-124) for evidence against an exceptionless reading of "whenever" in the above quote.

¹⁰As is apparent from a brief section about distinguishing motives from other types of disposition (Ryle, 1949, p. 110), Ryle is not even very whole-hearted about sharply delimiting the category.

concerning Ryle's conception of the modal force of disposition ascriptions, and his methods of individuating the conditions and doings which single out motives-as-such, seem inessential to the previously introduced core of his view. I mention them only to put them aside.

Having done so, we may note how well the good parts of Ryle's disposition-
alism about motives seem to dovetail with a further but brief pronouncement made by the later Davidson:

If a person is constituted in such a way that if he believes that by acting in a certain way he will crush a snail he has a tendency to act in that way, then in this respect he differs from most other people, and this difference will help explain why he acts as he does. The special fact about how he is constituted is one of his causal powers, a disposition to act under specified conditions in specific ways. Such a disposition is what I mean by a pro-attitude. (Davidson, 2004, p. 108)

Now viewing Davidson's pro-attitudes as of a piece with the other types of attitude we've been considering, we can see Davidson as saying that, not only is citing a desire to explain an action citing a disposition to do what is thereby done, but, more specifically, that it is citing a disposition to do the type of action at issue in conditions where the agent believes that what's done is a way to do that type of action.

Davidson gives these explanations a shape that must be constitutive of action if my account of the nature of action is correct: If acting is exercising a desire to do a type of action in conditions where there's a belief which represents the agent as doing that type of action, and if citing such a disposition is explanatory of an agent's action, then the applicability of the form of explanation which Davidson apparently describes must be constitutive of action.

If someone has indeterminate doubts about the cogency of such explanations, these may be alleviated by noting that the form of explanation is familiar from other places: Just as it is explanatory to appeal to a reflex, or a property of brittleness, or a photosynthetic propensity, it is explanatory to appeal to a desire, since all are dispositions, mentioning which functions to give a certain kind of information about why the subjects at issue would do what they are doing. Someone might, I suppose, query this proposal further, by asking whether, even in these other cases, we have genuine examples of explanation. Perhaps this complaint would rest on the idea that ascriptions of dispositions are not, as such, explanatory. But that is a very large and implausible assumption, which I will set aside.

But what about the more local complaint that, whether or not mentioning dispositions can in general be explanatory, it could not be explanatory to explain someone's action by appeal to a disposition called a desire, if the presence of a desire is entailed by doing the action? Analogously, one might question whether brittleness can explain breaking if breaking is construed in such a way that it entails brittleness. I will not get into any such analogous worry about the explanatory force of entailed dispositions. But below I will examine this *type* of objection, which essentially questions the idea of a dispositional and yet constitutive form of explanation, as it pertains, specifically, to the case of action.

4.3.3 Nagellian Objections

In this spirit, Nagel has at one point sought to undermine the thesis that “a desire of the agent must always be operative if the action is to be genuinely his” (Nagel, 1970, p. 27). But the discussion which follows this statement of Nagel's target suggests that Nagel is really out to undermine two thoughts, both of which might be described by something like the above quote, though they are

really quite distinct.

The first thought which Nagel seems out to undermine is that the *reason* for an agent's action is always a desire on the part of this agent. At least that is the thesis which I take Nagel's target to be when he suggests, for example, that where someone acts for a reason, "the presence of a desire is [not] a necessary condition for the *presence* of the reason" (Nagel, 1970, p. 30).

The second thought which Nagel seems out to undermine — albeit perhaps it is only an intermediate aim for him, is that an agent's action is always explainable by appeal to the agent's desire. At least that is how I interpret his apparent claim that, where someone acts for a reason, "[having a desire to do it] is a necessary condition of [the] efficacy [of their reason], but only a logically necessary condition. It is not necessary either as a contributing influence, or as a causal condition." (Nagel, 1970, p. 30).

On the surface level, at least, these two objections seem distinct. For it seems perfectly in order to say that every action is explainable by a desire, even if not every action is done on the basis of a reason that is a desire. And in fact that is precisely the line I will take. But since Nagel makes the two objections in the same pages, and doesn't emphasise their distinctness, it can seem that someone who goes along with Nagel's first objection must go along also with his second one, either because the first objection somehow entails the second, or because it derives its force from the second. But below I want to show that neither is the case.

For Nagel's objection to the idea that an agent's reason is always constituted by this agent's desire seems to have the following shape:

N1 Subjects act, and when they do it is necessary that they want to do the performed action (Nagel, 1970, p. 29).

Though Nagel does not explicitly say why he thinks this, this is a reasonable

and, as I have argued, true claim.

N2 When someone acts, their desire to do the performed action is either based on a reason, or it is a desire which has “simply assail[ed]” them (Nagel, 1970, p. 29).

The contrast here is one between desires formed by thinking about the reasons there are for wanting things, and desires like hunger, which tend to have a more arational aetiology. I do not know whether Nagel thinks desires in the second category need be especially passionate or brute, but it does not seem necessary to think so for making his argument. Something like a sudden desire to doodle, touch an elbow, or make a funny remark might fit into the same category as suddenly impinging hunger.

Nagel then suggests this:

N3 But it is not the case that, whenever someone acts, their desire to do the performed action has simply assailed them (Nagel, 1970, p. 29).

Here it seems that Nagel thinks that, if we were not to accept N3, we would have to endorse the absurd claim that doing something wanted is always like acting on a sudden urge or craving, or (I suppose) like acting on a sudden impulse to touch an elbow or make a funny remark, making impossible “motivational action at a distance” (Nagel, 1970, p. 27).

Nagel then seems to say the following:

N4 When someone acts, and their desire to do the performed action is based on a reason, then this reason is not always (and perhaps it is never) the presence of a further desire (Nagel, 1970, p. 30).

At least, says Nagel, there is “no reason to believe” that every desire is adopted on the basis of another desire, and seems much more reasonable to think that

sometimes (or often, or maybe always), something like a consideration concerning future happiness, or the welfare of others, is the reason on the basis of which a desire is adopted (Nagel, 1970, p. 29).

Nagel then takes this further step:

N5 When someone acts, and their desire to do the performed action is based on a reason that is not the presence of a further desire, the presence of this desire is not this subject's reason *for this action*.

Nagel fills in this last claim by saying that “[o]ften the desires which an agent [...] experiences in acting will be motivated exactly as the action is”, and even motivated, to make the point more explicit, “by precisely what motivates the action” (Nagel, 1970, p. 31). He suggests that, in such cases, we should not say that this desire is the reason for this action, but rather that the reason for the desire is equally a reason for that action. For an example of this, we might say that, although it might be entailed by someone's helping their neighbour that they want to, the reason for both the helping and the wanting to might be something that is not a desire, such as, perhaps, the neighbour's need for help. And of course a single example like this is really all that Nagel needs to show that someone can act on a reason that is not a desire, nor in any important sense constituted by the presence of their desire.

Nagel's argument proceeds by introducing a series of choices (is every desire unmotivated or not? and must a motivated desire be adopted on the basis of a further desire or not?) and arguing, in each case, for the adoption of one choice, finally arriving at N5, and thereby rejecting views which try to construe an agent's reasons as always constituted by this agent's desires. I find each of the five claims, including N5, entirely plausible. But it is quite clear that none of these claims introduce the idea that a desire is somehow incapable of explaining an action.

Still it is true that Nagel does make some such claim. And it might be thought that the claim does in some implicit way support his argument. For it can seem that a claim about the explanatory ineptitude of desires is supposed to provide some support to N4. To bring out how it might be thought to support it, consider first a possible opponent to Nagel, proposing that whenever a desire is based on something — something we may perhaps call a “reason” for this desire — this in fact amounts to the desire being based on a further desire. Such an opponent would be committed to supposing that forming, say, a desire to help someone that needs help, is really a matter of realising some fact about one’s own desires, and adopting some further desire on the basis of that one: “Wait, I want to help people who need help — so I want to help this person, who I realise does need help.”

This kind of view is misguided for making practical reasoning seem oddly insular, but Nagel’s argument, in so far as I have presented it up to now, does not clearly explain what is misguided about it. It can then seem that some claim about the explanatory meekness of desire is meant to perform this duty, and thereby support N4. But if so it is very unclear how that is supposed to work. It might be thought that when Nagel says that desire “is not necessary either as a contributing influence, or as a causal condition” (Nagel, 1970, p. 30) when someone acts, he simply means that desires cannot explain actions at all. He could then make the following very simple sort of argument: Desires can’t explain actions, but reasons can, so reasons aren’t desires. But if Nagel thought that this was a good argument, it would be unclear why he goes through the trouble of providing the above five-stage argument, which is quite involved.

In any case, the argument does not seem to be a very good one, since the suggestion that desires are incapable of explaining actions is both implausible and unargued. Hence, in sum, Nagel’s argument, against desires-as-reasons

can be read in two ways, neither of which makes it plausible that desires can't explain. On one reading, the argument, successful or not, leaves it open whether desires can explain. On another interpretation, the argument rests on the idea that desires can't explain, but does not really give grounds for thinking so.

To drive home how one might cling to the idea that desires can explain without thinking of desires as an agent's reasons, consider also Dancy's (2000, especially chapter 5) kindred argument, which, it seems, reinforces Nagel's conclusion about reasons, without forcing any particular view of the role of desires in explaining actions. This argument makes use of the idea of acting for a good reason, and suggests, quite convincingly, that someone prepared to postulate psychological items in the role of reason could not make sense of this idea. Dancy's argument, further, targets not only the thesis that an agent's reasons are always this agent's desires, but also other forms of "psychologism" about reasons. The core of Dancy's argument is, I think, that psychologism is the weakest member of this inconsistent triad:¹¹

Psychologism Whenever someone acts for a reason, their reason is that they believe or desire something

Normative Constraint Sometimes when someone acts for a reason, their reason is a good reason reason for them to so act

¹¹I acknowledge that this formulation misses out on much of the subtlety of Dancy's discussion. For example, it ignores his preceding discussion of the relative weaknesses of some varieties of psychologism, opting rather to tweak his main argument in such a way that it clearly targets all. It also reconstitutes an opaque claim, on which Dancy puts a lot of emphasis, that "motivating and normative reasons should be capable of being identical" (Dancy, 2000, p. 106) as the claim, below, that agents sometimes act for reasons that are good. The intent of the two formulations is probably the same, but Dancy's formulation has me thinking of a set of motivating reasons, a set of normative reasons, and an intersection between them, which seems to require some principled conception of what goes into the respective sets, inviting metaphysical detours about the nature of reasons. Worse, it seems to require a prior conception of a motivating-or-normative reason which makes such intersections possible, inviting a charge of begging the question. Dancy seems alert to this in noting that his argument "rests an enormous amount on" (Dancy, 2000, p. 106) the emphasised claim. I am hoping to provide a slight dialectical improvement, below, by presenting Dancy's normative constraint in such a way that it does not *introduce* anything like his conclusion, but rather introduces platitudes which we have no clear way of making sense of without abandoning psychologism about reasons.

Anti-insularity Sometimes when someone acts for a reason that is a good reason for them to so act, their reason is not that they believe or desire something

The two first claims undermine the third, since if reasons were always believings or desirings, and some reasons for an undertaken action are good reasons to so act, no good reasons for an undertaken action could be anything but its agent's believing or desiring something. Dancy rightly calls the resulting view "implausible" and "extreme" (Dancy, 2000, p. 100).¹²

Someone endorsing psychologism is then only left with the possibility of denying what Dancy calls the "normative constraint". But, though Dancy (2000, p. 106) doesn't quite put it in those words, that is just implausible in its own right, for reasons that have nothing specifically to do with psychologism. Why be so sceptical as to think there is no such thing as acting for a good reason? Consequently we should conclude that psychologism is to be denied. And I see nothing wrong with the conclusion.

But as suggested, my issue with Dancy's argument is not with its soundness but with its scope. As many (including Darwall, 2003; Millar, 2004; Davis, 2005; Setiya, 2007; Smith, 2012; Alvarez, 2009, 2010) have noted on encountering the conclusion of Dancy's argument, the thesis Dancy targets seems quite readily shed (and perhaps it should never have been taken on) by those who want to provide an account of acting, or acting for a reason, which ineliminably mentions its explanatory connection with desire. The resulting dialectical situation seems to be this: Even if Nagel and Dancy are right to say that an agent's reason for their action need not be their desire, the theorists who are nominally targeted by their arguments are *still* free to say that an agent needs to want to do what they are acting in doing, in a sense which, for all Dancy shows by

¹²In fact it is hard to so much as make sense of this kind of insularity, since it is not so clear we can coherently think of an agent who views all their reasons as facts about their own psychology.

the aforementioned argument, is constitutively explanatory of their action. If there are arguments which put this kind of “psychologism” about acting to rest, they are not instances of the present sort of argument, which only concerns psychologism about reasons.¹³

Having put reasons aside, it seems that a further sort of argument or contention can be discerned in the alluded-to passages of Nagel’s writing, which is not specifically about the nature of the reasons for which people act, but is rather meant to threaten the idea that an acting subject is constitutively subject to a mode of explanation which mentions desire. The rough and ready way of understanding this Nagellian contention is as follows: If the presence of desire is guaranteed when someone acts, then mention of desire can’t amount to a genuine *explanation* of why someone acts. There is something intuitively compelling about this neglected argument. How could a description explain an action if what it says about the acting subject is entailed by the subject’s acting? But in fact this argument has a sound version with a benign conclusion and an unsound version with a malign one. Below is the sound argument:

Consequence If N is acting in Aing, it follows from this that N wants to A

Non-Explanatory (C) If N is acting in Aing, mentioning that N wants to A
does not explain that N is acting in Aing

To avoid a muddle, note that we should not be encouraged to dismiss this Nagellian argument on the basis of Davidson’s famous and swift dismissals of various kinds of “logical connection argument” (Davidson, 2001a, p. 13-15). These

¹³Dancy (2004) gives a sceptical response to the idea that two kinds of explanation are applicable when a subject is acting for a reason. But I do not really understand the reasons for this scepticism. Much of Dancy’s criticism seems to hinge on the idea that since a causal explanation citing beliefs and desires would be available even where there was no appropriate “fit” between these states and the action explained, the applicability of such a purely causal explanation could not be constitutive of action. But this objection seems to rest on a too narrow construal of causalism, sensible versions of which should require some such “fit”. In any case, Dancy explicitly leaves room, near the end of the article, for a *dispositional* form of explanation which is constitutive of action. As it happens, that is just the sort of explanation which I am claiming to be constitutive of action.

hinge on the idea that one event may be the cause of another (for example, that Jones lost weight because on hunger strike) even though we may refer to the resulting event in a way which implicates its cause (for example, I suppose, by giving a description like “Jones lost the kind of weight that could only be lost by going on hunger strike”). In such an example we would have entailment from one description to another (from “Jones lost the kind of weight that could only be lost by going on hunger strike” to “Jones was on hunger strike”) although items constituting the truth of the two descriptions still entered into causal relationships (Jones’s hunger strike *was*, of course, the cause of losing whatever amount of weight thereby lost). Hence entailment from one description to another does not preclude the items constituting the truth of the respective descriptions from entering into causal relationships.¹⁴

But however good Davidson’s objection is deemed to be, it does not give us a convincing way of undermining Nagel’s quite different sort of “logical connection argument”, which concerns *explanation*. An odd statement like “Jones lost the kind of weight that could only be lost through hunger strike because... Jones was on hunger strike!” is hardly a worthwhile explanation if “...” divides explanans from explanandum, and this is not undermined by Davidson’s observation that logical relationships don’t preclude causal ones, since one can clearly mention a cause of something without explaining it. And though some may miss this, the original comments Davidson makes in connection with dismissing logical connection arguments do not seem so much as intended to defend the possibility that something, such as that someone acted in doing something, may be explained by something which it entails, such as that they wanted to do it. Davidson *does* say that something may be explained by something which

¹⁴Relating to an issue raised in footnote 20 on page 32 and footnote 3 on page 37, this Davidsonian observation can perhaps be adapted so as to be acceptable even to someone who does not view causation as a relation between events considered as particulars, but as a relationship between things “more closely tied to the descriptions of events” (Davidson, 2001a, p. 150), such as, for example, that Jones lost weight and that Jones was on hunger strike.

it is entailed *by*, when he suggests that “It dissolved” can be explained by “It’s water-soluble and was placed in water” (thinking the former to be entailed by the latter, thus dismissing the whole idea of a “dispositional masker”). But he never so much as suggests that something along the lines of “it dissolved in water” could be explained by “it’s water-soluble” where the latter is thought to be entailed by the former. Here he opts, instead, to suggest that there is *no* such entailment (Davidson, 2001a, pp. 14-15), which of course simply evades the present line of inquiry, which concerns whether, if there is such entailment, it precludes explanation.

The conclusion of this excursus is only this: Davidson and others have shown that a pair of descriptions can stand in entailment relations, although the objects or facts which make these descriptions true can enter into causal relationships. But this does nothing to prove that a description which entails another can explain it. The ordinary Davidsonian wisdom about logical connection arguments is thus powerless to undermine the simple and compelling Nagellian argument which was given above.

Further, I can see nothing wrong with the conclusion of this Nagellian argument, and think we should grant it. For what sort of insight, information or elucidation could be provided by an explanation whose explanans was a straight consequence of its explanandum? Someone might be impressed enough with the structural features and possible truth of such a string of descriptions to want to call it an explanation, but that would seem to make for an unhappy notion of explanation, which intuitively would render it quite unimpressive to say that action is constitutively subject to a mode of explanation which mentions desire.

We should note, however, that the above argument does not destroy every hope of saying that action is constitutively subject to a form of explanation which mentions desire. It only warns against a naive way of understanding that

thesis. And there seems to be nothing about *Consequence* which defends that more general and more serious claim, as can be seen by considering the below crude argument:

Consequence If N is acting in Aing, it follows from this that N wants to A

Non-Explanatory 2 (C) If N is acting in Aing, mentioning desires on part of N does not explain that N is acting in Aing

This argument is clearly not valid, and its conclusion is clearly false. Even if someone's desire to A can't figure in a real explanation of their acting in Aing, talk about what people want *is* capable of figuring in informative explanations of action, as when someone explains someone's going to the bank by appeal to their desire to repay a debt.

But, in itself, this doesn't defend constitutivism about desire explanations, since there is nothing obviously constitutive about an explanation of someone's action of going to the bank which merely mentions their wanting to repay a debt. As is likely to be noted, someone can perform such an action without such a desire (perhaps they are there to make a deposit). Someone pushing Nagel's line of argument might be tempted to conclude from this that, when descriptions are entailed, they are not explanatory, whereas when they are explanatory, they are not entailed, and try to make use of this fact in an argument against constitutive action explanations which mention desire.

And I think there would be *something* right in such an objection, although it misunderstands and therefore misses its target. For if someone is acting in going to the bank, this does not entail the applicability of any particular description, apart from ones equivalent to "the subject wants to go to the bank". And descriptions equivalent to "the subject wants to go to the bank" are not explanatory of someone's doing an action of going to the bank. Saying that the subject wants, say, *to repay a debt* might be explanatory, but of course it is not

entailed, and the same thing might be said about wanting to make a deposit, wanting to steal stationery, or whatever else might explain someone's going to the bank. So it seems true that no particular description is both entailed and explanatory when someone is acting in going to the bank.

But, to remind ourselves, *Constitutive Action Explanation* does not require that any particular explanation is entailed. It requires only that there's a *form* of description, mentioning a *kind* of explanans, such that an explanation of that form is possible whenever someone acts. Since the above specimens are all of a form such that they explain an agent's action by mentioning their desire, *Constitutive Action Explanation* would be safe if, though no particular desire ascription was entailed by someone's going to the bank (except their wanting to go to the bank), *some* desire ascription, like "they want to repay a debt", or "they want to see how people behave in long lines" could always explain someone's going to the bank.

How, then, might such a thing be guaranteed? There are, after all, apparently some things which a subject "just wants" to do, with no *further* desires of the sort that might explain the action. Given that, in such cases, no such further desires can figure in true explanations of the subject's action, and given also that, by previous reasoning, a simple desire to do what one wants can't figure in an *explanatory* description of why the action happened, it might now seem impossible that there should be an all-encompassing notion of constitutive action explanations which mention desire. But this complaint rests on a conflation of two forms of explanation. For, although it is true that there is not always a further desire in the sense of a desire to achieve a further thing, so that the action might be explained by appeal to a further end, there must, I think, always be a further desire in the sense of a more specifically individuated desire — a desire for a narrower type of outcome than that the pertinent action

is performed.

Perhaps it is not necessary that a subject that is acting in going to a bank has a desire to do something else thereby, such as repay a debt, but it does seem necessary that such a subject has a more determinate desire, embodying a more determinate grasp of a more determinate type of bank-going outcome, such as, for example, a desire to go by foot or by train, or to do it slowly with a knife in hand. I propose, further, that we can capture this further specificity, always present where a subject wants to do a type of action, by using the now-familiar connective “in order to”:

Explanatory Whenever N is acting in Aing, there is a true explanation of the form “N is acting in Aing because N wants to X in order to Y”¹⁵

Explanatory describes a form of explanation which mentions a specific form of explanans, but this form of explanation can be instantiated in two ways. Either it cites a desire to take a particular means to the action at issue, or a desire to take the action at issue towards a particular end. Hence, either the “A” above corresponds to the “X”, or to the “Y”, with some other action taking up the other position in either case, thereby yielding information about how the agent represents their action and (to however a limited degree) how this agent might be liable to act on other occasions where similar ends or means might be available. And this thesis is not seriously challenged by the above considerations, since even in those cases where an agent has no further end in

¹⁵It might seem that the possibility of wanting to do something in order to do something else is not supported by the account of wanting I favour. But the idea of a desire to do something should, as per the suggestion in chapter 2.3.2 on page 65, be identified with being disposed to act in doing that. Hence wanting to turn left is having a disposition to do an action of turning left. It might still seem, I suppose, as if the present account doesn’t countenance the idea that someone can want (or be disposed) to do something *in order to* do something else. But that is a natural elaboration of the account I favour: For if every action necessitates the presence of a desire to do this type of action, and if someone can do a whole course of action of the type “turning left in order to buy some nails”, then they seemingly must have a desire to turn left in order to buy more nails, the exercise of which is performing this entire course of action.

view, so that their action can't be explained by appeal to a desire for such a further end, the agent must, at least, have some means in view in acting, so that their action *can* be explained by appeal to a desire to take a further means.¹⁶

And it seems clear that, if a subject is acting in picking flowers, mentioning that they want to do so in order to put them under their pillow gives information of why that action is happening. It is true, as Nagel insists, that just on the basis of knowing a subject is acting in picking flowers, we know that they want to do so. But we do not know the more specific nature of this desire, which we may then learn of in such an explanation. Once we are told that it is a desire to pick flowers in order to put them under the subject's pillow, we know that the subject is disposed to do what it takes to act in picking flowers in a set of situations more restricted than those involving encounters with flowers. For then we know of this subject's more specific liability to pick flowers when they deem that doing so can contribute to putting them under their pillow.

What may seem somewhat less clear is that mention of an agent's wanting, say, to get their knife in order to pick flowers, is informative about their picking of flowers. But I believe this is a mistake, which again rests on conflating explanations by desires for means with explanations by means. Someone's getting a knife in order to pick flowers doesn't explain their picking flowers, except perhaps in the limited sense of telling someone how they managed to do it. But someone's wanting to get a knife in order to pick flowers does explain their picking flowers, since it tells us that the agent wanted to get their knife, and, further, that they thought that was an available way of achieving the wanted end of picking flowers. It thus gives us information about the agent, which sheds light on why it should have happened that flowers were picked.

¹⁶Given the above discussion of basic action, such an explanation must always be available, but even if it were deemed that basic actions (and so basic desires for means) are needed, *Explanatory* would not fall down, since all it requires is that every action comes with a desire for a mean *or* an end — not that every action comes with a desire for a means, or that every desire comes with a desire for an end.

4.4 Desire Ascription

4.4.1 An Equivocal Puzzle

Suppose that saying something like “I want to catch a fish” or “this person wants to have lunch” is generally a matter of ascribing to the mentioned subject a desire with a satisfaction condition that corresponds to the verb phrase employed. If we suppose this, then saying “I want to catch a fish” or “they want to have lunch”, ought in general to be a matter of ascribing, respectively, a desire with someone’s catching a fish, or someone’s having lunch, as its satisfaction condition.

But now consider the everyday phenomenon, recently attended to by Fara (2013) and Lycan (2012), that someone can generally say something like “I want to catch a fish”, although if they should end up catching a small or otherwise unexpectedly disappointing kind of fish, it would seem false, or at least misleadingly incomplete, to say that they thereby did what they said they wanted, or that they satisfied their ascribed desire.

In just what way does the everyday phenomenon clash with the view we just supposed? Below I want to show that, as Fara and Lycan construe these cases, they commit to denying the supposition that is at issue, but that, in so doing, Fara generates insurmountable obstacles to providing any plausible account of the meanings of desire ascriptions, whereas Lycan is forced into a mysterious and troublesome account of what it takes to satisfy a desire.

I will then suggest a different way of understanding the everyday phenomenon, on which it does not falsify the original supposition, but instead highlights a neglected feature of what it is to want to do something, and a correspondingly neglected way in which desire satisfaction can be partial or truncated. I hope to explain both features by appeal to the idea that desires can be instrumentally complex.

But to make clear why Fara and Lycan have bad responses to the felt tension

between this supposition and the everyday phenomenon, and to make clear why my response is a good response to this tension, I will first need to make explicit just what the supposition of the first paragraph comes to.

4.4.2 On Desire Ascription

To make an explicit view out of the original supposition, while simultaneously defending its innocence, I want to show how it is a straightforward consequence of a pair of theses which are themselves innocent. One concerns the semantics of desire ascription, and the other the metaphysics of desire satisfaction. There is, I think, nothing very contentious or new in the view I will present, nor in the thought that it rests on a combination of metaphysical and semantic commitments.¹⁷ But that is the point, since soon I want to make it explicit that although Fara and Lycan officially target more contentious doctrines, they become equally committed to denying something that is close to common sense.

The semantic leg of the upcoming thesis goes as follows:

Semantic Platitude A statement of the form *N wants to A* is generally used to express that the subject mentioned in place of “N” wants to do the kind of thing mentioned in place of “A”.

This thesis should not be controversial. It only has such benign consequences as, for example, that if “Elizabeth wants to brush her teeth” is used as such phrases generally are used, it expresses that Elizabeth wants to brush her teeth. And of course the example is chosen at random. It will give analogous consequences for any statement of the form *N wants to A*.

¹⁷Braun, writing on the same topic, briefly outlines a view that is somewhat similar to the view I will describe, calling it “the plausible view” (Braun, 2015, p. 141), signalling that it is commonly held. Anticipating what I say here, Braun also notes that this view results from a combination of semantic and metaphysical claims (Braun, 2015, p. 143). Something like the ordinary view I will describe also seems implicit in much of the literature that views desires as relations to propositions, and in the literature which views desire ascriptions as describing subjects as standing in relations to propositions.

It is important to note that it would be a mistake to replace “generally” with “always”, or “necessarily” since it clearly is possible, and does happen, that someone uses a statement of the form *N wants to A* to express something which is not this subject’s wanting to do the mentioned kind of thing, or (perhaps more often) their wanting something which is not *precisely* the mentioned kind of thing. For example, someone could say “Elizabeth wants to brush her teeth”, using the whole string of sounds as code for something totally different, or they could be using “wants” ironically, or say a statement like that although “Elizabeth” was a slip of the tongue, where the intended subject was really someone named Elias Beet. As will be more pertinent to the subsequent discussion, someone could also say “Elizabeth wants to brush her teeth” and thereby express that Elizabeth wants to brush only her upper teeth, or that Elizabeth wants to brush a prime number of teeth, or any such variation on the claim that Elizabeth wants to do *something* which is a case of her brushing her teeth. The point of “generally” is that it allows such possibilities, so that they do not undermine the semantic platitude.

Perhaps it will now be asked what “generally” is supposed to mean. But I will not try to answer this difficult question. I will note that the subsequent discussion is neutral on whether “generally” is taken to mean something like “most of the time”, or, as is perhaps more plausible, “paradigmatically” or, or even, perhaps, something like “when it expresses the proposition that is expressed by salient other statements of the same form, including statements made in languages other than English”. And I will note that, however exactly it is to be construed, “generally” cannot be dispensed with, as long as, on the one hand, there is some unity between all the various uses of statements of the form *N wants to A*, whereas, on the other hand, divergent or non-standard uses are possible.

The second claim, perhaps not platitudinous outside of philosophy, but reasonably platitudinous within it, is that a desire to A is the sort of attitude that would be satisfied just in case its subject should end up having succeeded in Aing. One extremely common route to some such conclusion goes via viewing desire as a propositional attitude. It involves saying, first, that for N to want to A is for N to stand in a relation to the proposition that N succeeds in Aing, and, second, that this proposition is or corresponds to the satisfaction condition of that desire (Searle, 1983; Crane, 2001). But this route to the conclusion is apparently not safe, as there are various objections to propositionalism about desire, several of which (I submit) don't have ready responses (Brewer, 2006; Thagard, 2006; Montague, 2007; Thompson, 2008; Merricks, 2009). In the next two paragraphs I describe what I take to be an especially troublesome issue with propositionalism about desire, which I have not seen a satisfactory propositionalist response to:

Thompson invites a more serious look at what it's supposed to mean to say that a desire to brush one's teeth, or go to school, is really an attitude with the object *that N brushes their teeth*, or *that N goes to school* (Thompson, 2008, pp. 127-128, including footnote 11).¹⁸ He argues that, once we do thoroughly consider such propositional adaptations of ordinary desire ascriptions, we'll discover that they assign the wrong meanings to these expressions. For the only available ways of understanding such propositional ascriptions says Thompson, are these: The habitual one, on which "N wants that N walks to school" becomes equivalent to "N wants that N (is the sort of person that) *walks* to school", the progressive one, on which it amounts to the same as "N wants

¹⁸It is often remarked that there is some linguistic discomfort to saying that a desire to brush one's teeth is, as it were, really an attitude towards the outcome *that one brushes one's teeth*. But after being noticed, this discomfort is often ignored in the hope that such wrinkles can be dealt with (Crane, 2001, pp. 111-112), or overruled on the grounds that there is nothing for a desire to be, or for a desire ascription to ascribe, but a propositional attitude (Searle, 1983, pp. 29-30). Thompson can be taken as showing the conceptual source of the linguistic discomfort.

it to be the case that N *is walking* to school”, and, finally, the perfective one, on which it amounts to something like “N wants that N *has walked* to school”. The trouble Thompson finds with propositionalism is that none of the three available readings captures what seems to be ascribed by saying that N wants to walk to school: The habitual reading is clearly off, the progressive one seems compatible with indifference about reaching the school, whereas the perfective seems compatible with indifference about active participation in getting there.

I suspect Thompson is right that none of these three readings of “N wants that N walks to school” can work. But it is a tall order to prove that there could not be another, more felicitous propositional reading. Could we not, for example, simply introduce a special verb phrase in the relevant propositions, signifying *both* active participation *and* ultimate success, such as, for example, “going through with” the undertaking, or “doing it from start to finish”? Allowing this, we could read “N wants to walk to school” as something along the lines of “N wants that N *goes through with* walking to school”. This still looks like an ascription of a propositional attitude, and seems to avoid the troubles of the earlier proposals. But it could be objected that “going through with walking to school” is still not quite what is wanted where someone wants to walk to school, or that it is merely a cumbersome way of sticking with the letter of propositionalism while rejecting its spirit.

I do not want to make more of the present section about this objection — that would be a long and tangential discussion — but fortunately this does not seem necessary, since it seems possible to capture the spirit of the propositionalist’s rationale for the near-platitude without using wording that forces us to view the object of desire as somehow propositional. For we may say simply this: For N to want to A is for N to have a desire with Aing as its object. For Aing to be the object of N’s desire is for this desire to be such as to become satisfied

just in case N ends up having A-ed. This more neutral rationale generates the following thesis:

Metaphysical Near-Platitude For a subject, N, to want to do a kind of thing, A, is for N to have a desire which will be satisfied just in case N ends up having A-ed.

By way of illustration, this thesis has consequences like the following: For Elizabeth to want to brush all her teeth just is for her to have a desire of the sort that will be satisfied by her brushing all her teeth. (Again, it will have analogous consequences for any other example of wanting.)

The phrase “ends up having A-ed” is necessary to get over a hurdle about time, but one which seems distinct from present concerns, and perhaps also distinct from the just-bracketed concerns about propositionalism. The hurdle is that if someone wants to bake a cake, and if they baked one yesterday, so that it is true that they have baked a cake, this does not mean that their present desire to bake a cake is satisfied. What we need to do justice to is that someone only satisfies their desire if “the future take[s] some specific course” (Fara, 2003, p. 147). But the idiomatic “ends up having A-ed” already accommodates this, since on reflection this must come to something like “ends up having A-ed in the future as seen from the present (by the one with the desire)”. It doesn’t mean the same thing as “will have A-ed in the future”.¹⁹

Taken together, the platitude and the near-platitude straightforwardly entail the following view, which must, then, be true if they are:

¹⁹There might be thought to be the further issue that, if someone wants to do something but succeeds in doing it only after giving up on wanting to do it, their desire should not count as satisfied. But I suspect that it should. If someone wants to smash a ball-like object that is flying towards them, and realise too late that it is an egg, so that they no longer want to smash it as they do, then I think it would be right to say that their desire to smash it is satisfied, only no longer present. Such a person succeeds in doing what they *wanted*, albeit not in doing what they want. Perhaps I am wrong, and a subject may only satisfy their desires if they still have them. But if that is true, it could easily be incorporated into some amended version of the metaphysical near-platitude, without causing any trouble for the subsequent discussion, which does not concern the satisfaction of abandoned desires.

Desire Ascription A statement of the form *N wants to A* is generally used to express something which is *true just in case* the subject mentioned in place of “N” has a desire which will be satisfied just in case they end up having done the kind of thing mentioned in place of “A”.

And on initial appearance, this view only seems to have consequences that are benign. It tells us, for example, that if someone says “N wants to turn left for Kilimanjaro”, and uses the statement along general lines, and what they say is true, then N has a desire which will be satisfied just in case N ends up having turned left for Kilimanjaro. And it seems to tell us such things as that, wherever there is a desire of the sort that would be satisfied by (and only by) doing some kind of thing, like turning left, this desire can be felicitously ascribed by saying the subject in question wants to turn left.

4.4.3 Semantic Revisionism

If someone says they want to walk towards school, then could they satisfy the ascribed desire by doing so at a pace slow enough to miss the last lesson? Or, suppose, instead, that someone is said to want to eat cake. Could the talked-about desire be satisfied by this person’s eating a microscopic piece of cake?

However exactly we are to respond to such questions, we should note that they are generally applicable, so that it seems that we should come up with a general strategy for answering them. In fact it seems that, whenever we think of a desire ascription, we are able to think of an outcome where the subject in question ends up having done the thing they’ve been said to want to do, although at the same time it can seem false, or at least somewhat misleading or incomplete, to say that this subject has thereby satisfied the ascribed desire.

Fara and Lycan have come up with some cases like the above, and both seem to say that, in these sorts of cases, a desire to do something is ascribed,

although doing the thing mentioned in this ascription is not enough to satisfy that desire. Fara is most explicit on this point:

[W]hen Fiona *says* that she wants to catch a fish, she does not express a desire that becomes satisfied just in case the proposition that she catches a fish becomes true. That proposition becomes true if she catches a tiny minnow. But her desire does not thereby become satisfied. Moreover—and this is the point I wish to emphasize—her self-ascription of the desire is, despite all this, true. The desire that makes her claim true has a more specific content than the proposition expressed by her embedded clause. (Fara, 2013, p. 254)

Fara’s official target is not *Desire Ascription*, but a related thesis, which seems to come to this: A statement of the form *N wants to A* expresses that N stands in a *relation* to the *proposition* that N As, and that N stands in that relation to that proposition just in case that desire is satisfied in precisely those *possible worlds* in which N As (Fara, 2013, pp. 252-254).²⁰

This related thesis seems quite close to *Desire Ascription* — seems to be a version of it which adds a particular brand of propositionalism to it, thereby encountering the aforementioned problems about tense, while ignoring the issue of futurity which I quote Fara herself as raising in the previous section. But here I will not consider how Fara’s characterisation of the puzzle case affects this related thesis, since I think that thesis is a more specific but less plausible version of *Desire Ascription*. I want to consider how Fara’s characterisation of the puzzle cases affects *Desire Ascription*.²¹

²⁰Hence I mostly agree with Braun’s (2015, p. 147) interpretation of what Fara targets.

²¹The theory Fara targets “seems to be a more specific and more detailed version of the plausible view”, notes Braun (2015, p. 147), and on this I agree. Braun does not seem to consider the possibility that, in targeting the more specific and more detailed view, Fara might undermine something less doctrinal and closer to common sense, but only notes that Fara’s objections, if successful, “throw[...] some doubt onto the [less specific] view” (Braun, 2015, p. 161).

Fara's characterisation of the case of Fiona involves these two claims: That Fiona speaks truly when she says "I want to catch a fish", and that Fiona does not have a desire of the sort that would be satisfied just in case she should end up catching a fish (Fara, 2013, pp. 255-256). Hence Fara endorses, or is committed to endorsing, the following pair of theses about Fiona's case:

Truth Fiona's statement of "I want to catch a fish" is true

Opaque Ascription Fiona's statement of "I want to catch a fish" is not made true by the fact that Fiona has a desire such as would be satisfied just in case she should end up catching a fish

Of course this pair of theses entail that Fiona's statement of "I want to catch a fish" does not express the presence of a desire of the sort that would be satisfied just in case Fiona should end up catching a fish. But this, taken on its own, does not commit Fara to rejecting *Desire Ascription*, since there remains the possibility that Fiona's statement is somehow used in a deviant or non-standard way.

But Fara does not seem to think there is anything non-standard about Fiona's statement, since in her discussion she takes this puzzle case to undermine a thesis concerning "the relation expressed by 'wants'" (Fara, 2013, p. 250), meaning, presumably, the relation that is in general or paradigmatically expressed by that word, and assuming, presumably, that the word is used in that general or paradigmatic way in Fiona's case.

More importantly, the assumption that Fiona is using the statement along general lines seems perfectly in order, and hard to undermine. For there is no apparent reason to say that the imagined Fiona is using the words "I want to catch a fish" as code for something else, or that she's non-overtly specifying some further aspect of the object of her desire (as someone would if they said "I want to ride in a *car*", with special emphasis, and perhaps pointing towards

an especially fancy specimen).²²

So even though it is not explicit in Fara's writings, it seems that, on top of the preceding, Fara should think the following:

Usualness Fiona's statement of "I want to catch a fish" is used as statements of the form *N wants to A* are generally used

This trio of theses, two of which Fara clearly is committed to, and one of which it seems that anyone should be committed to, straightforwardly contradict *Desire Ascription*. For if all the theses are true, then in Fiona's case, a statement of the form *N wants to A* is used as such statements generally are used, and this statement is true, although this statement is *not* made true by the fact that Fiona has a desire of the sort that would be satisfied just in case she should end up having done what is mentioned in place of A in her statement. And then this statement could not express that Fiona has such a desire.

But now I want to argue that, in characterising this case in such a way that it falsifies *Desire Ascription*, Fara generates a problem which it seems she can't get out of. I want to do this by raising the question of what Fiona's statement could plausibly express, if not the presence of a correspondingly easy-to-satisfy desire. And I will argue that there is nothing that it could plausibly express, if not this.

Fara's way of answering the question seems to begin with the assumption that the psychological fact which makes Fiona's statement *true* is that "Fiona wants to catch a fish that's big enough to make a meal" (Fara, 2013, p. 250). Because it is extensional, this does not tell us what Fiona's statement expresses, though of course it sets constraints for what it could express. But the answer is also gappy in a further respect, since it doesn't tell us what, in Fara's view,

²²If for some reason it should be thought that there is something deviant about Fiona's ascription, it is also worth reiterating that it seems that Fara's concern with this ascription should affect just about any ascription to the same extent — we can always think of a viscerally unsatisfactory way in which someone can do what they've been truly said to want.

it means to say that Fiona wants to catch a fish that is big enough to make a meal. The only thing that Fara clearly seems to say about this is that it is having a desire which requires for its satisfaction *at least* catching a fish that is meal-sized.²³

So Fara seems to think the following:

Underspecification Fiona’s statement of “I want to catch a fish” is made true by the fact that Fiona has a desire which requires something over and above Fiona’s catching of a fish for its satisfaction (including that the fish is meal-sized, and perhaps also that it is not rotten, that it is alive, and so on)

Whatever Fara’s own motivations might be for taking on such a claim, and in spite of the fact that I will reject other elements of Fara’s view, it is worth noting that *Underspecification* is an inherently attractive thesis. If Fiona has a desire of the sort that requires for its satisfaction that she ends up catching a fish that is above some size (and perhaps also not rotten or foul-tasting, and perhaps indefinite further conditions need to be added here), it still seems right that, in saying “I want to catch a fish”, Fiona is, albeit in an incomplete way, describing *this* desire. And this even though Fiona’s statement does not seem — at least not overtly — to represent anything about size, freshness, or flavour. Hence Fara seems right to say that an account of what such a statement means must “explain[...] how a desire report [like Fiona’s] could be true even when its

²³It is not so clear whether Fara thinks, for example, that ending up catching a meal-sized fish is *sufficient* for satisfying Fiona’s desire. Fara never says she subscribes to anything like our *Metaphysical Near-Platitude*. If she would endorse it, then she would think that wanting to catch a meal-sized fish is to have a desire of the sort that would be satisfied *just in case* one catches a meal-sized fish. Then Fara would think that, although in saying “I want to catch a fish”, Fiona was not referring to a desire that could be satisfied by catching a tiny minnow, she *was* referring to a desire that could be satisfied by catching, say, a dead, rotten fish. This proposal makes Fiona psychologically odd, since now she is taken to care about whether a caught fish is meal-sized, but not about whether it is otherwise meal-appropriate. Since this odd supposition does not seem necessary for Fara’s purposes, and since I do not think anything in the case compels us to think of Fiona in this way, I do not ascribe such a commitment to Fara here.

embedded clause underspecifies the desire that makes the report true” (Fara, 2013, p. 267).

But back to the question of what Fiona’s statement could mean if not that Fiona has a desire of the sort that would be satisfied just in case Fiona should end up catching a fish. If we adopt Fara’s assumption that the statement is made true *not* by the fact that Fiona has a correspondingly easy-to-satisfy desire, but by the fact that Fiona has a harder-to-satisfy desire, then we must say that Fiona’s statement means something which is consistent with it having truth conditions as follows:

Opaque Ascription Fiona’s statement of “I want to catch a fish” is not made true by the fact that Fiona has a desire such as would be satisfied just in case she should end up catching a fish

Underspecification Fiona’s statement of “I want to catch a fish” is made true by the fact that Fiona has a desire which requires something over and above Fiona’s catching of a fish for its satisfaction (including that the fish is meal-sized, and perhaps also that it is not rotten, that it is alive, and so on)

In trying to say what Fiona’s statement means while making it consistent with these desiderata, we’ll run into insurmountable problems. For a start, consider the proposal that we simply add whatever extra conditions are supposedly required for satisfying Fiona’s desire to the meaning of her statement. Suppose, that is, that in saying “I want to catch a fish”, Fiona just is saying that she has a desire which requires some extra conditions, over and above her catching a fish, for its satisfaction. That interpretation of her statement would, of course, straightforwardly conform to the above pair of theses. But it is a non-starter, for at least three reasons, all having to do with its over-specificity:

The first problem is that, if Fiona’s statement meant such a thing, then

someone would need to know that Fiona's desire wouldn't be satisfied with catching a tiny minnow in order to know what Fiona meant by her utterance. But it clearly is possible for a hearer to understand what Fiona is telling them, without knowing that. (Imagine someone is told by Fiona that she wants to catch a fish, but that they remain unsure whether Fiona is fishing for sport, or to make a meal, or for relaxation, or for thrills, or to catch the tiniest fish possible to use for bait later).

The second problem is that, if Fiona meant to convey some such narrow desire by her statement, then it is not clear how someone interested in catching a somewhat different type of fish could felicitously respond "so do I". But of course they could. (Imagine that someone is told by Fiona that she wants to catch a fish, and this someone is interested in catching a tiny minnow, and wants to give expression to their shared interest).

The third and perhaps most straightforward problem is that intuitively Fiona's statement would stay true if Fiona's desires should change and start to target some different size of fish, or a fish of no particular size. (Imagine that Fiona changes her mind about meal-sizedness, and wants, instead, to catch a smaller fish, or just any fish. Her original statement still intuitively applies.)

What this shows is that it must be some quite general fact about Fiona which gets expressed by her statement — something which can be known without knowing too much about Fiona's interests, and truly co-ascribed without sharing too many of Fiona's interests, and retained by Fiona without retaining too many of her further interests. This observation can seem to motivate a sort of proposal with which Fara experiments, on which Fiona's statement is to be interpreted by introducing open variables. Perhaps Fiona's statement means, for example, that she has a desire of the sort that would be satisfied only by catching a fish of some particular type, where the type in question is not conveyed by her

statement (Fara, 2013, pp. 257-258)?

If we take Fiona's statement to express the thought that Fiona has a desire of the sort that would require the catching of some particular type of fish, then Fiona's statement will plausibly be true where Fiona has the more specific desire, which required meal-sizedness for satisfaction, but also true if Fiona's preferences should change, as in the case where she changes her mind and starts to be interested in catching a tiny fish instead. It would also arguably make sense of how Fiona's desire ascription could fit another subject with different size-preferences, and of how someone could know what Fiona meant without knowing too much about her size-preferences.

But such a proposal could still not work, since, if Fiona should grow totally indifferent to which type of fish she ends up catching, her original statement would still intuitively be true of her, although it would not be true of Fiona that she had a desire of the sort that requires for its satisfaction that she catch a fish of some particular type. Even with an open type variable, the proposal still assigns a too specific meaning to Fiona's statement, merely in virtue of introducing the idea that her desire is type-specific.

If it does not seem that conditions concerning types of fish figure in the meaning of Fiona's statement, it might still seem that there are *some* conditions, over and above the catching of a fish, which do figure in it. And it might be thought that we could provide an adequately general account of the meaning of her statement by saying that it expresses that Fiona has a desire of the sort that would be satisfied only if she should catch a fish *in some or other condition*, albeit — still — a condition which is not conveyed by her statement.

Although Fara never considers it, this interpretation of Fiona's statement might be thought, as the previous one, to have the right truth conditions in the limited sense that catching a meal-sized fish can be described as catching a

fish in a particular kind of condition — one where the caught fish is meal-sized. It also would accommodate that Fiona’s statement could stay true in a future where she becomes indifferent to which type of fish she catches. (As an added bonus, it seems to be at least compatible with the idea that, when Fiona says she wants to catch a fish, she expresses the presence of a desire of a sort that would require for its satisfaction that the caught fish is not dead, rotten, or useless in a number of other ways.)

But is it entailed by Fiona’s statement, as we intuitively understand its meaning, that she has a desire of the sort that would be satisfied only if she should end up catching a fish *in some condition*? For of course if that is not *entailed* by Fiona’s statement, it could not *mean* that. In answering this question, we must be careful. For if we read “Fiona catches a fish in some condition” as co-entailed by “Fiona catches a fish” (thinking that, surely, anything that happens happens in *some condition*), then there will be no clear obstacle to saying that Fiona’s ascribed desire is of the sort that will be satisfied just in case she catches a fish. So this way of reading Fiona’s statement makes it conform to *Underspecification*. But it does so at the cost of contradicting *Opaque Ascription*. For on the present sort of reading, Fiona’s statement attributes to her a desire of the sort that would be satisfied just in case Fiona should end up catching a fish. So, even if the present reading is correct, *Fara’s project*, which we took as an effort to conform to both theses, will be undermined if that is what Fiona’s statement means.

But as soon as we read “Fiona catches a fish in some condition” as not co-entailed by “Fiona catches a fish”, but as requiring some more specificity from Fiona’s desire than that statement, it will no longer clearly be true whenever Fiona’s statement seems to be. For imagine that this statement requires, additionally, that Fiona’s desire is targeting something somewhat more particular

than the mere catching of a fish. Whatever extra condition is imagined here, it seems possible for Fiona to decide that she does not care about attaining that extra condition, although her original statement still applies. For example, we might take the extra requirement to be that the weather is not terrible, that the fish is not inedible or dangerous, or that things do not take too much time, or perhaps that things go as Fiona imagines they should go. Whatever extra condition we come up with, it seems possible for Fiona's desire to stay true in conditions where her desire changes so as to not target such conditions. For example, Fiona's statement will intuitively stay true even if all of a sudden she starts to strive for an opposed outcome where she catches a terrible, useless fish, of the kind that is sure to bite her finger when she takes it off the hook, as she's standing in the rain and counting her disappointments.

In sum it seems that, while accepting *Underspecification*, the only plausible interpretation we could find of Fiona's statement was one which contradicted *Opaque Ascription*. As soon as we tried to add anything to the meaning of this statement about any further condition required for satisfying her desire, we seemingly ended up with false interpretations, since it always seemed possible for Fiona's statement to stay true even when her desires stopped targeting such proposed conditions. This should make us ask why, with Fara, we should accept *Opaque Ascription*, which seems to preclude the only acceptable interpretation of Fiona's statement which we could find. Is it really as plausible, as Fara took for granted at the start of her article, to deny that Fiona's statement of "I want to catch a fish" is made true by the fact that Fiona has a desire such as would be satisfied just in case she should end up catching a fish?

4.4.4 Metaphysical Revisionism

But someone of broadly the same persuasion as Fara might now respond by trying to backtrack out of the present difficulties, by denying not *Opaque Ascription*, but *Underspecification*. The idea in such a proposal would be to say that Fiona’s statement of “I want to catch a fish” is *not* made true by the fact that Fiona has a desire such as would be satisfied just in case she should end up catching a fish, *although*, at the same time, that Fiona’s statement was not in the business of saying that Fiona’s desire requires anything further for its satisfaction — not, for example, in the business of talking about the type of fish — even de re — which Fiona would need to catch to satisfy her desire.

In his discussion of closely related matters, Lycan seems to end up painting some such picture. For drawing inspiration from McDaniel and Bradley (2008), Lycan attempts to analyse the desires of analogous puzzle as follows:

[The kind of desire present in the puzzle cases is] a relation, not between a subject and a proposition, but between the subject and two propositions: such a desire has both an “object” proposition and the relevant “condition.”

The condition may be tacit and quite complex; it need not be fully represented, [by the subject with the desire, nor by a statement ascribing it].

If nearly every desire ascription has a vague and messy “condition” parameter, then finally it is straightforward that our problematic desire-ascriptions are correct but so are the protests. You want lunch, on the condition that the proffered lunch is decent and pleasant. You want fame, on the condition that it be based on something admirable. Absent those conditions, your desire is not satisfied, even though it is semantically satisfied. Your assertion, “I want to have

lunch,” was true in the context, and so was your protest that you did not want to have the lunch I force-fed you. (Lycan, 2012, p. 209)

According to Lycan, desire is not as simple an attitude as is suggested by our aforementioned metaphysical platitude. Ignoring the issue of whether the object of a desire is properly thought of as a proposition, the quotes suggest that, according to Lycan, puzzle desires require for their satisfaction *both* the attainment of an associated object *and* compliance with an associated “condition parameter”. If puzzle desires have these two components, then it becomes possible, I suppose, to say that the statements in puzzle cases are describing only one of these components — the “object” of the pertinent desire — but that these ascriptions are not thereby saying anything much at all about the *satisfaction condition* of that desire.

And Lycan’s suggestion seems to be that this is just what is happening in our puzzle cases: Someone says someone wants to do something, and what they say is straightforwardly made true because it corresponds to the *object* of the desire, with no hidden parameters or variables, but still the pertinent desire is not such that it will be satisfied just in case this object is attained. For *satisfaction* of that desire requires compliance with a condition that is distinct from attainment of its object — which was “not fully represented” in the statement. Fiona says she wants to catch a fish, and thereby straightforwardly ascribes to herself a desire with the “object” that she ends up catching a fish. It is just that satisfaction is not so easy to come by as attainment of a desire’s object is. In Fiona’s case, satisfaction might require also that the fish is meal-sized, and perhaps a number of further things.

From a logical point of view, Lycan’s proposal gives us a way to conform to *Opaque Ascription* without also conforming to *Underspecification*, thereby avoiding the troubles that seemed to affect anyone with Fara’s commitments

(as does, of course, a proposal on which Fiona’s statement is made true by some random mathematical fact, or that Fiona wants something, or whatever else is true). But the cost of the proposal is that it introduces a notion of a desire’s “object” which is not so readily understood. In what sense is Fiona’s desire “directed at” catching a fish, if catching a fish is not the satisfaction condition of her desire? And on the other side of things, what could it be for a desire to be satisfied, if not merely for its object to be attained?²⁴ It is not clear that these notions, normally tightly intertwined, have any real content once we try to pull them apart. In any case, Lycan does little to show how to make sense of the resulting pair of notions.

4.4.5 Deepening the Ordinary View

We should now take stock. We’ve discovered that, if we accept both *Opaque Ascription* and *Underspecification*, and say that statements of the form “N wants to A” are not in general made true by the presence of desires that are as easy to satisfy as these statements overtly suggest, then we won’t be able to assign any plausible meaning to these statements, since the assigned meanings will either have truth conditions that are more demanding than the intuitive truth conditions of these statements, or, if not, will undermine *Opaque Ascription*.

We then considered Lycan’s more mysterious proposal, which affirmed *Opaque Ascription* but did away with *Underspecification*, by suggesting that the statements in puzzle cases specify objects, but not thereby satisfaction conditions, of asso-

²⁴The only real information Lycan provides on this matter is in the form of a hint to the effect that the satisfaction conditions of our desires must be fixed by facts about our brains, and perhaps by Fodorian representations somehow in these brains (Lycan, 2012, p. 211). But I suspect that this answer could not be filled in further, since we simply have no clear hold of the idea of a desire’s satisfaction condition which is not provided by viewing that condition as that desire’s object. If — somehow — it could be shown that Fodorian representations in a subject’s brain determine the type of outcome that would satisfy their desire, then it would be overwhelmingly tempting to say that this was a representation of the object of that desire as well. Not only that, it is not clear what could be said against this, to try to pry objects apart from satisfaction conditions, since it is not clear what independent criterion of a desire’s object there could be.

ciated desires. That seemed to buy us out of the previous dilemma at the large cost of making it very unclear what an object or a satisfaction condition of a desire might be.

These discoveries make it overwhelmingly tempting to reject *Opaque Ascription*, while maintaining *Desire Ascription*. Having done so, there will be no reason to seriously doubt our *Semantic Platitude* and our *Metaphysical Near-Platitude*, as no reason to doubt these theses has been presented which is distinct from the insistence on *Opaque Ascription*. Because of this, it will no longer be motivated to entertain the idea that what it is to want to A can come apart from what it is to have a desire such as could be satisfied just in case one should go on to A, nor the idea that, in standard cases, saying statements of the form *N wants to A* can refer to anything but the mentioned subject's wanting to do what is mentioned in place of A. Hence below I will take for granted that *Desire Ascription* is true, thereby also taking for granted, for example, that having a desire that is satisfied just in case one ends up having A-ed can be described as wanting to A.

But to make this conclusion satisfying, something needs to be said to accommodate the pressures which leads these authors to introduce *Opaque Ascription*. What, first of all, is this pressure? To give a very provisional expression of it, something has surely been left out if we declare that Fiona and the rest satisfy their ascribed desires in the bad outcomes mentioned, and leave things at that. Or, to approach the thought from a different angle, there surely is some crucial difference between someone who says they want to catch a fish, and catches a tiny minnow, and someone who says they want to catch a fish, and catches a reasonable fish of the sort that they had in mind to catch. What is the difference?

Part of the answer, I think, is in *Underspecification*. I earlier noted that this

thesis has independent plausibility, which does not derive from *Opaque Ascription*. The appeal is that as long as we are assuming that Fiona does not have a very unusual and manic sort of desire *merely to catch a fish*, it seems right to say that when Fiona says she wants to catch a fish, she is somehow referring to a further desire, which would require more for its satisfaction than simply her catching a fish. We want to say things like this: “Sure Fiona wants to catch a fish — that really is an outcome that she is going for. But that isn’t all there is to the desire she is talking about — she doesn’t *merely* want to catch a fish.”

But the puzzle this generates is one of reconciling our endorsement of *Underspecification* with our denial of *Opaque Ascription*. How could it be that Fiona, for example, is self-ascribing a less specific desire, *and at the same time* incompletely describing a more specific desire, merely by saying “I want to catch a fish”? To make the problem explicit, consider what we now seem committed to, while abstracting, this time, from Fiona’s particular case:

Transparency in General In puzzle cases, a statement of the form *N wants to A* is made true by the fact that the subject mentioned in place of N wants to do what is mentioned in place of A

Underspecification in General In puzzle cases, a statement of the form *N wants to A* is made true by the fact that the subject mentioned in place of N wants to do something over and above what is mentioned in place of A

Since it seems that there is not necessarily any “speech act pluralism” happening in the puzzle cases, it must, moreover be that, in these cases, a single statement, with a single meaning, manages simultaneously to require the existence of both sorts of desire for its truth. But how could a pair of desires be metaphysically entangled in the way they would need to be for such a thing to be possible? Building up to a full explanation of what is happening in the puzzle cases, I

want to make sense of such metaphysical entanglement by appeal to the following thesis:²⁵

Desire Divisibility Part of what it is for N want to A in order to B, is for N to want to A, and want to B

To see what this thesis means, start by noting that it must be possible to have a desire with the object that one does something in order to do something else.²⁶ For example, someone can have a desire to do a whole project of the type “pouring tea from a pot in order to fill a cup with it”. *Desire Divisibility* is the claim that whenever there is a desire like this, with the object that something is done in order to do something else, the subject with the desire wants both to take the means and to achieve the end.

People might object: Can’t I want to get on a plane in order to get to Brazil, but dread planes? This kind of objection is a mistake. Wanting to A is compatible with all kinds of negative attitudes towards Aing, and I think that very little supports the present kind of objection apart from some confusion between wanting and what we might call “liking”, or perhaps between wanting and “wanting + liking”. The mistake is exposed by the fact that, just as I can want to get on a plane without liking the idea, I can want to go to Brazil without liking the idea, or even want to get on a plane in order to go to Brazil without liking the idea. Whatever wanting to do something is, it does not entail enjoying the idea of achieving the kind of outcome wanted.

²⁵The material that follows is indebted to Anscombe (1963), and also to Thompson (2008). It is an attempt to carry over some of their insights about instrumental action to an account of instrumental desire. I allow that, in my account of desire, I’ll end up diverging from what Anscombe says about wanting. In the case of Thompson the difference is more elusive. Thompson seems to essentially identify wanting to A with being in some early stage of Aing intentionally. I resist this identification. But much of what I say about my kind of desire seems applicable also to what Thompson says about his kind of desire. Both exhibit means-end structure that mirrors (in my case) or simply is (in Thompson’s case) the kind of means-end structure that their objects — actions — exhibit.

²⁶See note 15 on page 135 for some discussion of how the notion of a desire as a disposition to act, together with the idea that actions can come in instrumental complexes, supports the idea that a desire can take an instrumental complex as its object.

And I think it is clear that any other objection to *Desire Divisibility* trades on introducing analogous illicit demands on desiring. These demands are illicit in part because they seem to threaten not just that subjects want to A and B when they want to A in order to B, but also that subjects want to do any of the things we seem to think they want to do. For example, one might try the objection that a subject can want to A in order to B without wanting to B, because Bing doesn't "occur" to the agent. I want to be nice to my family members in order to express that I care about them, but perhaps I don't think about expressing that. But of course it need not occur to me that I want to be nice to my family members either, in the relevant sense of "occurring", nor does it seem like most of the things a person wants need to "occur" in the mind of the person in that sense.

Analogously, someone could say that I can want to turn a crank in order to grind black pepper without it being salient to me that I'm interested in turning the crank, or that I will not be disposed to point it out if someone asks me what I want. But, just like the previous objection, these ones exclude too many things from counting as objects of a subject's desires, for none of these conditions seem to generally apply when people want to do things. This is a clue to the insight that wanting isn't intimately associated with emotions, occurrent thoughts, salience, conversational dispositions, or conscious awareness.

But what is the argument *in favour* of *Desire Divisibility*? I want to make it plausible, and shed some light on it too, by identifying a quite bloodless and inconspicuous attitude which a subject must have towards the relata in the in order to-relation, guaranteed to be present wherever a desire to A in order to B is present, and then identifying this attitude as one of desiring to do these things. The argument is not a straight deductive argument for *Desire Divisibility* that flows from independent premises, but this is to be expected, if (as Kant famously

seemed to think about a similar claim) *Desire Divisibility* is an analytic truth about desires.

That a subject wants to A in order to B seems equivalent to the subject wanting to take the *means* of Aing to the *end* of Bing. Such a subject must then have some kind of attitude towards Aing, which we might call “wanting to do something *by* Aing”. And equally the subject must have some kind of attitude towards that end, which we might call “wanting to do something *in order to* B”. When I say that wanting to A in order to B entails wanting to A and wanting to B, I just mean that these two resultant attitudes *are* desires. And I do not know how that might be congruently denied. For if we think of these attitudes as not involving wanting the means or the end, we’ll always, I think, misconstrue these attitudes. Adopting A as a means to Bing is surely more than thinking of Aing as a possible way of Bing, and what could the difference be if not one of wanting to A? Likewise, taking Bing as an end of Aing is more than thinking of Bing as a possible result of Aing, but what could this difference be if not one of wanting to B?

Once we get this feature of desires in view, I think we’ll see that it is somewhat similar to, and no more controversial than, the feature of beliefs whereby, for instance, believing a conjunction entails believing each conjunct. If it’s worth calling this feature of beliefs an “analytic” truth about that kind of attitude, I think it’s worth calling *Desire Divisibility* by the same word. But ignoring analyticity, we may declare, a little more carefully, that *Desire Divisibility* is not a claim about which desires tend to be formed by people, or even which desires it would be rational for them to form. It really is a straightforwardly metaphysical claim about what it is for someone to want to do one thing in order to do another.

Desire Divisibility now affords us a way to show that, in the puzzle cases, a

statement like “I want to catch a fish” or “someone wants to wear high heels” can be, simultaneously, an ascription of a desire with a correspondingly simple satisfaction condition, and an underspecified description of a desire with a more complex satisfaction condition. For we may now suppose that things are as follows in the puzzle cases:

Instrumental Hypothesis In puzzle cases, the subject mentioned in place of N wants to do what is mentioned in place of A *in order to* do something further — B

This hypothesis clearly, neatly and generally explains how the ascriptions in puzzle cases may simultaneously fully specify a subject’s desire, and underspecify a further desire. If the hypothesis is correct, then Fara was correct in supposing that the ascriptions in puzzle cases underspecify the objects of associated desires. For these puzzle cases were then ones where someone wanted to take a means and thereby achieve an end, but where an ascription specified merely the means. And this is, I think, an exceedingly natural way of thinking of what is happening in these cases. For example, it seems right to say that the reason why Fiona’s ascription is partial or incomplete is that she wants to catch a fish *in order*, for example, to make a meal. Similarly, if a Heidi says she wants to wear high heels, but won’t be very satisfied with wearing them on her hands, a ready explanation is that she is really referring to a desire to wear high heels *in order to* look a certain way, or make a certain impression, or walk somewhere, or whatever.

It might be thought that this explanation is at best partial, because it might be thought that it is possible for Fiona to catch a fish, to make a meal out of it, and still not have satisfied her ascribed desire, because of some further badness about this outcome. But it is plausible that the true explanation of this further narrowness to such a desire is that there is *further* instrumental complexity to

it. If catching a fish and making a meal out of it still isn't quite satisfactory to Fiona, that is plausibly because Fiona wanted to catch a fish, in order to make a meal, in order to *eat it*, but ended up with something inedible, or because she wanted to make a meal in order to impress someone, but the dinner was not impressive, or something along such instrumental lines. So it still seems intuitive that the present account of the desires in the puzzle cases provides a full explanation of the narrowness of these desires, though ultimately this approach may require introducing quite long chains of instrumental desires: Fiona wants to catch a fish in order to make a meal in order to eat it in order to be healthy and taste something nice in order to ..., and so on.

It turns out to be very hard to think of a case where there is no such instrumental complexity, although the case at hand is clearly a puzzle case in of the sort that is described by *Neatness-in-General* and *Underspecification-in-General*. Try to imagine, for example, that Fiona *just* wants to catch a fish, but not in order to do anything further. If this is really what Fiona wants, then it is hard to make sense of the suggestion that somehow, the ascription "I want to catch a fish" is an *underspecification* of any of her desires. If Fiona does not want to catch a fish in order to do something, then it seems right to say that she really just wants to catch a fish, in the simplest sense of that statement that we can imagine. Or try to imagine that Fiona, for example, *just* wants to catch a meal-sized fish, although all she says is that she wants to catch a fish. Now it no longer seems plausible that, in so doing, Fiona ascribes to herself a desire merely to catch a fish, which again precludes the case from being a puzzle case in the sense considered here.

Desires can take complexes of means and ends as their objects, where the means and ends are wanted in virtue of wanting the complex. Once we see this, we see that there is no need to set up a dilemma, as Fara and Lycan

implicitly seemed to do, between saying that ordinary desire ascriptions are first approximations to some deeper psychological truth about a subject, and saying that these ascriptions wear psychological truths on their sleeves. For if desires are complex in the way I suggest, both such claims are straightforwardly shown to be true together.

Chapter 5

Circles

5.1 Introduction

In this chapter I consider objections to accounts which say that acting requires a belief about acting. I consider three kinds of objection to the substance of this account, and one more troublesome worry about the shape of the account.

First, I consider charges of chauvinism: Could not animals and small children, for example, act without believing it? I argue that, properly developed, such charges are either deeply inconclusive, because they do not say why we could not attribute such notions to such subjects, or simply question-begging, because in trying to make it an empirical question whether such subjects have these notions, they already suppose that acting without such notions is possible.

Second, I consider an objection to the effect that the present kind of account makes for an infinite series of beliefs about beliefs whenever someone is acting, and thereby places an impossible psychological demand on an acting subject. I respond to such objections by appeal, among among other things, to the idea that infinite nested beliefs about acting need not make for correspondingly

infinite demands on the psychology of the acting subject.

Third, I consider an objection to the effect that, because the present kind of belief requires adoption for truth, it violates some notion of what it is for a belief to be objectively true, so that on the account no one could believe, or objectively believe, that they're acting. I respond to this objection by rejecting those conceptions of objective truth which seem to underpin it.

After these objections, I consider a different kind of objection which, although it admits that the account might be true, questions its informativeness, by suggesting that there is an inherent and unavoidable opacity to it. In reply to this, I adopt a hint from Anscombe, and argue that, since believing one is doing an action is believing one is doing another action in order to do it, no question about what it is to believe one is doing an action need ever be left unanswered.

5.2 Objections to Reflexive Belief

5.2.1 Chauvinism Charges

The view I have taken on acting might seem to generate commitments that are objectionable because chauvinistic. Here I lay out two such commitments, and consider two types of chauvinism objection, extracted from some work of Mele's.

In deriving the targeted commitments, consider that I claimed that for N to act in Aing requires N to be in possession of a bit of psychology which represents N *as* acting in Aing. It does not seem far-fetched to infer from this that for N to act in Aing requires N to understand *what it is to act in Aing*. Name this the first consequence of my account. Going further, if the original claim is taken together with this consequence of it, they seem to generate a further

consequence. For assuming that acting in Aing requires understanding what that comes to, and, as we started out saying, that part of what it comes to is possession of the aforementioned bit of psychology, it is hard to avoid inferring that for N to act in Aing requires N to believe *that N is in possession of a bit of psychology required for acting*. Name this the second consequence.

It might seem desirable, for someone with a certain set of philosophical instincts, to accept the original claim while circumventing one or both consequences. Such a theorist would need to introduce some way of blocking the first or second inference, and, to make a really satisfying case, some way of showing — as it were — that there's some smudge on an ordinary human's introspective lens which makes it possible for us to believe that we're acting, although without understanding just what that comes to, or, alternatively, to understand that we're acting, without believing that this involves our believing it. To the extent that such a theorist would be successful in maintaining my sort of view without going along with one or both consequences, they would apparently evade the charges raised below, since these seem to target these consequences more than my original stance. But I don't want to take this evasive line since, to me, the inferences of the previous paragraph seem in order, and the two consequences unobjectionable.

I should confess I have some trouble understanding just what such a semi-transparency view would come to (hence the slightly discouraging metaphor involving smudge on a lens in lieu of a clear statement of how semi-transparency might arise). To circumvent the first consequence, one might try invoking a vague or hazy kind of belief, such that someone can believe that p without understanding just what p comes to (thereby explaining how people can believe they are acting without understanding just what it is for them to act). To circumvent the second inference, one might try introducing materials from

the literatures on Frege puzzles and *de re* belief (thereby — somehow — allowing people to count as believing that they’re acting *without* believing some necessary consequences of that). I see no compelling reason to convince such a theorist to buy into my view so that they might see the need to respond to Mele’s upcoming objections. On the other hand, if my response below should be deemed appropriate, so that there is no deep problem with systematically attributing action-theoretic and psychological knowledge to acting subjects, these semi-transparency approaches may appear correspondingly devalued.

Mele (1999) criticises what looks like the first consequence in a brief discussion of Wilson (1989) and Ginet (1990). The central disagreement between Mele and these authors seems to be that they attribute to acting subjects a conception of what it is to act, whereas Mele thinks that, given certain facts about the “conceptual resources” of some agents, we cannot in general demand such understanding from an acting subject.¹ This makes it look like Mele’s argument is a straightforward appeal to counterexample: Not all agents must understand what acting is, since some agents could not. But in explaining why the set of agents outstrips the set of things that understand what acting is, Mele only really says that it seems “implausible” (Mele, 1999, p. 429) because “any notion of intentional action” attributable, for example, to children, must be “pretty thin” (Mele, 1999, p. 425), which just looks like a vaguer restatement of the negative thesis at issue.

Mele’s (1987) criticism of what looks like the second consequence has more to it, and this criticism can be adapted so as to affect also the first consequence. It emerges in his discussion of a thesis that is, at least, closely related to the second consequence, seemingly endorsed by Harman (1976, p. 441) and Searle (1983,

¹Mele also says that we can only attribute a bit of psychology representing the agent as acting as *the agent conceives* of acting — where this conception might be mistaken or incomplete (Mele, 1999, p. 427). But the purpose of this replacement doctrine is unclear to me. Perhaps he wants to preserve something he regards as an insight in the criticised proposals.

p. 85). For Harman and Searle have said that it is a condition on N's acting in Aing that N has an intention with the object that *this very* intention causes N to A.² Mele's objection starts familiarly and innocently enough by noting that "[a]gents capable of intentional action are beings with cognitive capacities and limitations[who] can intend only what their cognitive capacities permit them to intend" (Mele, 1987, p. 315). It proceeds by suggesting that representing intentions is more likely to lie beyond a subject's cognitive limitations than acting is, so that it would be reckless, in some epistemic sense, to make the former a necessary condition for the latter:

It is doubtful that very young children, dogs, and many other beings capable of intentional action [...] have the cognitive capacity to intend that an intention of theirs be executed. When my eight-month-old daughter intentionally crawls toward me upon seeing me enter the room, does she intend to crawl toward me by way of carrying out her pertinent intention[...]? (Mele, 1987, p. 315)

I suspect [...] not. But, in any case, we have not as yet encountered a convincing reason for endorsing the empirical thesis [...] that all intenders-to-A have a representation of the state or event that is their intention and of its resulting in their A-ing. And if the empirical thesis is not manifestly false, it is at least sufficiently uncertain that the safest course is to leave the matter open. In short, the very fact that the [theory] commits us to a stand on the matter is a strike against it - provided that there is a reasonable alternative that does not force us to come down on one side or the other of the empirical question. (Mele, 1987, p. 316)

²Here Searle apparently goes further than Harman, since his favoured condition is that the acting subject has an intention with the object that this subject "perform the action" of Aing (Searle, 1983, p. 85).

As is indicated by Mele's presentation, what he says here would be far from conclusive if taken as a straightforward argument from counterexample. Mele does not give us a compelling reason for denying that dogs and young children believe they are in whatever psychological states are required for them to act. But from how he elaborates on the matter, it seems that, though he grants Harman (1976, p. 441, n. 5) that "it is not clear what the test is for saying that the child has a concept of intention", he thinks that the test is at any rate empirical, and that it is, for some reason, implausible that agents will always pass it, so that it is unparsimonious or otherwise epistemically unsafe to go on supposing that the test may always be passed by an acting subject. I think, then, that Mele's argument has something like this shape:

M1 Dogs and young children are capable of acting

M2 There is only a slim empirical possibility that dogs and young children are capable of conceiving of whatever psychological facts are necessary for them to act

M3 (C) If an account makes it a requirement for acting that the acting subject conceives of the psychological facts which are necessary for this subject to act, this account is held hostage to a slim empirical possibility

If this exhibits the gist of Mele's objection, that objection has nothing specifically to do with intentions, nor with the possibility of "self-referential" intentions. The worry concerns the general thesis that the psychological features which figure in the true account of what it is to act need to figure also in the thought of the acting subject. What is supposed to threaten this thesis is not that there are straightforward examples of subjects who can act but can't have the requisite conceptions, but the empirical precariousness of supposing that there could be no such subjects. Having seen Mele's second objection in this light, we might

try to view the first one in the same light. It might then amount to something like the following argument:

M1 Dogs and young children are capable of acting

M2* There is only a slim empirical possibility that dogs and young children are capable of conceiving of what it is to act

M3* (C) If an account makes it a requirement for acting that the acting subject conceives of what it is to act, this account is held hostage to a slim empirical possibility

The second argument seems to be a generalised version of the first, since it says that an acting subject needs to understand what acting is, which, disregarding semi-transparency views, entails knowing about whatever psychological aspects acting has.

In assessing this pair of arguments, very much hinges on how we interpret M2 and M2*. And in assessing this pair of premises, we need to decide on a reading of “slim empirical possibility”. On a reasonably clear reading, the possibilities at issue are epistemic, so that the two key premises are saying that the capacities in question are not, *for all we know*, attributable to dogs and young children (or that we do not have *good reason* to believe they are, or that subjective *probabilities* that they are should be kept low).

The obvious problem with the argument, read in this way, is that the two key premises can easily be denied by those targeted by the argument. It would seem perfectly consistent, and, I think, reasonable, for Mele’s opponents to agree that dogs and young children can act, and to infer from this, in accordance with their view, that these subjects have the representational capacities in question, and to deny M2 and M2* on grounds of this conviction.

If, at this point, someone pushing Mele’s line of argument should insist

on preserving M2 and M2*, perhaps on the grounds that we have no certain knowledge about the contents of a dog's mind, it still seems perfectly in order for someone like Searle and Harman to accept these premises, but, on the grounds of this lack of conviction in the capacities of dogs and young children, to take back their endorsement of M1.

To illustrate the point, note that the arguments we're presently considering it have the following shape:

P1 p

P2 probably (for all we know) not-q

(C) probably (for all we know) not-(p→q)

Suppose that this form of argument is valid. It still is not clear how, on its own, such an argument could reasonably sway anyone to accept its conclusion. "Here is a potato", says someone. "That thing doesn't seem to contain any starch", says another. However convincing a case might be mounted for the second statement, it would not seem reasonable for the first person to infer from it that potatoes probably don't contain starch. For there remains the perfectly reasonable stance of taking back the original judgement: "If this doesn't contain any starch, I guess it is not a potato — perhaps it's a toy, or a different odd vegetable, then." Mele does not explain why the corresponding stance should not be open to someone who starts out thinking practical belief is definitive of action, but then becomes convinced to be a sceptic about about the reflective lives of dogs: If some very good case could be mounted for the case that dogs are probably too dumb to understand that they act, and, as a result, too dumb to understand that they possess the psychological features necessary for acting, then why should it not be supposed that dogs can't act, or probably can't?

Perhaps the reader will feel that the stance is less plausible in the present case

than in that involving potatoes. But all I am saying is that Mele's argument, on the present conception, makes no progress in proving that. There might be thought to be a more head-on approach of revising the arguments so as to introduce premises to the effect that, although dogs and young children can act, it would be possible to make the empirical discovery that they lack the representational capacities at issue. But as soon as this happens the arguments become question-begging. For it to be possible to empirically discover that a dog is incapable of conceiving of action or intention even though the dog is capable of acting *is* for there to be no deep connection between the former capacity and the latter of the sort which Mele's opponents suppose there to be.

Perhaps what really needs to be undermined is the general sympathy for Mele-style scepticism, which renders people all too ready to grant that, in principle, it would be possible for dogs and young children to act without having these capacities for thought about their actions. But since I don't fully understand the sources of this kind of scepticism, I do not know how to undermine it in any conclusive way. Awaiting a conclusive showdown, perhaps it will help to take note of the following simple sort of dilemma for people pushing such scepticism:

Very young children of some particular description either can or can't believe things (in the broadest sense) about what they're doing (in the broadest sense). If we decide that they can't, it starts to look mistaken to say that they *can* act in doing these things, and then they do not provide Mele with a case in point. But if we decide that such children *can* have beliefs about what they're in fact acting in doing, then what could be a better candidate for such a belief than one to the effect that what's done is due to the child in the way that actions are due to an agent?

5.2.2 Possibility Protest

On encountering the thesis that a certain kind of attitude is required for acting, which, further, represents its subject as acting, we may start to suspect that this thesis makes acting impossible, since the object of this attitude would have to include an infinite nested series of representations of representations of acting, rendering it “inaccessible to a finite act of thinking” (Kapitan, 1995).³

To give clearer voice to the suspicion, we should ask what it could mean to say that the amount of “complexity” of the supposed representations renders them “inaccessible” to ordinary subjects. And to do this, we first need to lay bare what this complexity is supposed to consist in. I think that it must be a kind of complexity introduced by the following argument:

Self-Awareness For N to be acting requires N to believe that N is acting

Transparency For N to believe that N is acting requires N to believe that whatever is necessary for N to be acting is the case

Infinite Self-Awareness (C) For N to be acting requires N to believe that N believes that N believes (and so on forever) that N is acting

Someone could, of course, try to stop the inference by denying *Transparency*. But as before, I know of no clear way of making sense of semi-transparency views concerning beliefs about acting. Hence rather than stopping the inference to

³There are analogous theses in other areas of philosophy, to which analogous objections might be made. An obvious example is the “KK principle”, on which knowing something requires knowing that one knows it. Though the query does not seem to have gained much prominence, one might ask, how, if infinite pieces of knowledge are required wherever there is knowledge, anyone could know anything. A recently more prominent objection targets a conception of knowledge how, on which knowledge how to do something requires contemplation of a proposition about how to do it, which — presumably — would require further contemplation of a proposition about how to do *that*, and so on. Stanley and Williamson (2001, p. 414) attribute an objection to this effect to Ryle (1949), and seek to undermine it. My upcoming response to the present thesis about action is, however, different from their response to the Rylean claim, but corresponds, roughly, to a response which they raise without pursuing (on the aforementioned page): That the episodes of contemplation and the propositions that are generated by the Rylean regress argument are “not distinct” from whatever episodes of contemplation and propositions this argument presupposes.

Infinite Self-Awareness, I want to ask why this thesis should make it impossible to act. Why is that kind of infinitely nested representation supposed to be out of reach to an ordinary subject?

It might be proposed that *Infinite Self-Awareness* requires someone to contemplate infinite propositions — one for every belief — and that this is impossible. But this seems misguided if it means that they must in some sense entertain the propositions. If, on the other hand, it just means that they must in some sense be “related” to each of these propositions, in the sense that any belief might be thought to require a relation to a proposition, then it is not clear why that should be thought problematic. Or it might be suggested that *Infinite Self-Awareness* requires infinite occurrent thoughts. But as was suggested in chapter 4.2 on page 112, there’s a mistake in requiring that the object of a belief must in this sense be present to mind. If, on the other hand, the proposal is merely that there are infinite beliefs, then it is not clear why that should place an impossible demand on a subject.

Hence it remains to be shown that *Infinite Self-Awareness* somehow generates an impossible psychological demand. On the basis of these examples it can rather seem that, where *Infinite Self-Awareness* generates an apparent infinity, this presents no real psychological difficulty for a subject, whereas, where someone thinks that there is some infinite psychological demand generated by infinite nested beliefs about acting, that demand, impossible as it might be, is not really generated by this thesis. None of this is to say that the notion of believing something does not place some demands or constraints on the believing subject. If someone believes something, they must not be dead, they must not be a tree, and they must perhaps have been causally in touch with the objects of the belief. (Very much more could be thought of here.) But there is no clear way in which any of these requirements — real as they are — could somehow

be infinitely multiplied by a thesis like *Infinite Self-Awareness*.

5.2.3 Objectivity Objection

A different sort of worry could be described as belonging to metaphysics. This worry is that, since we've characterised acting as involving a belief about acting, whereas it is a mark of objective truth that it obtains independently of human belief, it can never be objectively true that a subject is acting. Various versions of the independence premise will secure the worrying inference. The perhaps most attractively lean version figures in the following argument:⁴

Self-Awareness For N to be acting requires N to believe N is acting

Mind-Independence If some candidate for truth p is objectively true, p must be true independently of whether anyone believes that p

Irrealism (C) It is impossible for it to be objectively true that N is acting

Identifying truth as objective truth would lead to the further bind of making it impossible for it to be *true* that N is acting, but the argument is worrying enough as it stands, since even if non-objective truths are allowed, it does not seem a happy arrangement to put facts about acting in that category. And worrying as its conclusion is, the argument does seem to me perfectly valid. So

⁴There are other ways of construing objectivity as mind-independence, which do not specifically mention the truth at hand, but say, instead, that p is objectively true only if it is true regardless of whether *anyone* believes *anything*, or, yet more generally, that an objective truth is one which holds independently of whether anyone has *any attitude* concerning anything, or perhaps even independently of *any fact* about human subjects. These kinds of conceptions of objective truth threaten just about any account of action, but as has often been noted, they also undermine the very idea that there could be psychological truths (see Haack (1987, p. 281), Sober (1982, p. 371) and (Devitt, 1997, p. 16)), and hence such theses are seldom accepted in unqualified form. (Even eliminativists about psychology seem to prefer arguing for their position in less immediate ways than by simply bulldozing over it with the premise that objective truth obtains independently of human psychology — human psychology being, presumably, dependent on itself, so that no objective truths could obtain about it.) About the premise I have adopted below, it might be argued that it too precludes objective truths about beliefs, since it might be argued that if someone believes that p, they must believe that they believe that. I leave this question open, though I think the claim has some plausibility.

it will be no surprise that I respond to it by saying that *Mind-Independence*, at least in its fully unqualified form, should be deemed false.

Though it is difficult to decide on a standard by which to assess *Mind-Independence*, it seems that the point of endorsing it must be defending a kind of realism fit to make truth (as we want to say) non-arbitrary. But there is no clear sense of arbitrariness in which *Reflexivity* makes it an arbitrary question whether someone is acting in doing something. So whatever checks against arbitrariness realism is meant to secure, they seem achievable just as well by, for example, endorsing the thesis that an objective truth obtains independently of whether, *for some subject*, they believe it (as someone clearly *can* be acting in Aing independently of whether someone else thinks so), or else, in an admittedly less elegant way, by restricting the scope of *Mind-Independence* so that it concerns some restricted domain of reality — that which is in principle describable by a certain kind of non-psychologising science, perhaps.

5.3 Objection to Reflexive Accounts

In this section I return to the ponderous problem which we've encountered a few times, and especially at the end of the third chapter. The problem has been raised, and briefly discussed, by a smallish number of contemporary philosophers, (Ford, 2011; Haddock, 2012; Thompson, 2008, pp. 176-177, n. 14), but the sharpest formulation of it seems due to Anscombe:

We have the following situation about a type of concept. Let 'M' be a concept of such a type. [...] If an M-ing takes place, in that A M's, it is an essential constituent of the M-ing that A thinks (believes) he is M-ing. [...] Our trouble lay in the impossibility of explaining the *content* of the thought, and in the consequent impossibility of saying what it is for an M-ing to take place, not in any implication

that the content of the thought must in a way be conceived to be repeated *ad infinitum*. (Anscombe, 1969, pp. 64-65)

We are clearly in an Anscombian situation, since we have said acting is a type of concept instantiation of which requires the agent to believe they are acting. But what is the problem which Anscombe thinks this generates?

The problem seems distinct from all previously considered problems, which either targeted the realism of introducing such a situation, or the possibility of it. Anscombe does consider and dismiss one or two objections in the vicinity of these (Anscombe, 1969, p. 65), and does not seem concerned with the others. Her real worry seems to be that an account describing an Anscombian situation is useless not on pain of falsity, but on pain of not “saying” what its subject matter is at all. Though, as we’ll see, it is very difficult to adequately express this worry, it is also very easy to feel its force, by considering, for example, the following piece of philosophical dialogue: Someone asks what it is to act, and gets (as part of) the answer that it involves a belief about acting. Then they ask: What sort of belief is that? Since presumably the belief needs to be understood by specifying its object, and its object is acting, there’s a visceral sense that this interlocutor’s inquiry is going nowhere.

The reason why it is difficult to express this worry, of course, is that we have no clear way of articulating when a putative account ‘says’ or explains what its subject matter is, or as I’ll put it, when an account is informative. Hence our problem is not simply a problem of meeting some specified demand, but also one of specifying the demand in a tolerably clear way. Having noted this, we may think the ideal methodology would be to proceed with a crisp and general account of informative accounts against which to test various proposals for ways of “saying” or “explaining” what a practical belief is. But this supposedly ideal strategy does not seem available, as the following paragraph is meant to make

plausible.

There are very many examples of accounts which are obviously uninformative, of which we've mentioned some. "For N to be acting in Aing is for it to be true that N is acting in Aing" is one, "For N to be acting in Aing is for N to have a true belief that N is acting in Aing" is another. As was important for my argument in a previous chapter, "For N to be acting in Aing is for N's desire to cause N to A in a sense of 'cause' that is characteristic of *acting*" seemed uninformative too. Luckily there are also examples of accounts which seem informative. Aside from those reductive accounts which identify knowledge, science, love or colour through putatively independent concepts, there also seem to be accounts which *combine* circularity with informativeness. These include such accounts as specify a "state of knowing" as one standing in justificatory relations in a "space of reasons", where presumably what's at the other end of these relations is other states of knowing or knowledge-informed perceiving (Sellars, 1956; McDowell, 1996), or ones specifying beliefs as states figuring in internally coherent systems of further beliefs (Quine, 1960; Davidson, 2001b). These accounts seem informative, but why is an elusive question: They are no less circular than the uninformative specimens, they are equally liable to introduce semantic notions, and — contrary to a popular way of speaking of "good circularity" they seem to traffic in pretty "small" circles.

Here it might be suggested that one feature is shared by the informative but circular accounts. But I don't know how to describe it in a way that stands up to prolonged critical nudging. We might — vaguely and weakly — call the shared feature "holism". One account says that believing something is believing other things in a system of beliefs characterised by coherence. The other says that knowing something is knowing other things in a space of known things characterised by justificatory relations. But I don't know how to specify a clear

criterion for this feature of holism, and anyway I am not sure that it is the only way for an account to escape un informativeness. Still, it is worth noting that what we'll end up saying about action will have something like this feature, since part of the response to the charge of un informativeness will rely on the idea that an action exists in a context of further actions, as a belief about action must exist in a context of further instrumental beliefs.

Because the supposedly ideal strategy is not available, we will pursue a lower road below. Rather than beginning with a conception of informative accounts, and trying to show that our present account fits or can be configured to fit into that mould, we're going to begin with a test which exposes the kind of *uninformativeness* which Anscombe is worried by, and show how some of what we've already said about the nature of practical belief allows us to avoid that kind of un informativeness. The test is inspired by the feature of the imaginary piece of dialogue mentioned above, and goes as follows:

Test If there's an account of something, X, which itself mentions X, then if there's a question, "what is X?", which this account can only answer in a way which invites the same question again, then this account is uninformative, in the specific sense we're after.

This test does not say that an account is uninformative whenever it is circular, in the sense that its elucidandum figures in its elucidans. Rather, the test says that an account is uninformative whenever there's a question which it can only answer in a way which invites the same question again.

The former approach would pre-judge, in a quite incredible way, the question of whether there could be a benign form of circular account, declating futile the ambitions of various thinkers, including Strawson (1992, pp. 19-20), and, in an admittedly more remote philosophical setting, Dummett (1991, pp. 201-202), who have wanted to specify ways in which an account can be circular while

informative.⁵ It also would ignore Keefe's (2002, pp. 291-292) interesting view that, though benign circularity is possible, a strict account of it is not.⁶

Proceeding with our adopted test in the hope of yielding satisfying results, the first thing to note about it is that it can apply however many extra elements are introduced in a circular account. If someone says "acting is doing what figures in a belief that one is acting, and wanting to, and p and q", then their account provides whatever insight is provided by these extra conditions, but the difficulty about unformativeness remains at the heart of the account, as someone would still be in a position to ask: "What is the relevant belief about, over and above wanting to do something, and p and q?" Likewise, introducing practical commitments, takings as reasons, or even qualifying the kind of belief at issue by calling it a "causal" or "self-referential" or a presentation of an action under the aspect of the good, would still leave the same opacity at the centre of the account.

Noting this shows, I think, that a family of accounts of action or cognate notions (Harman, 1976; Searle, 1983; Velleman, 1989; Setiya, 2007), all set on introducing a kind of attitude with a kind of object which requires that kind of attitude, are uninformative despite their introduction of such extra elements. To see this, note, for example, what Searle says about acting in raising one's arm:

The [...] content of my intention [to raise my arm]must be at least
[...]that I perform the action of raising my arm by way of carrying

⁵It is true that these authors don't say enough to provide us with a clear criterion. Strawson's remarks on the matter are sketchy (constituting, I think, some version of the claim that largeness plus some other virtue can make a circle virtuous), and Dummett's are not clearly and directly applicable to our own case (involving, as it does, the idea that a logical law can be *relied on*, while not asserted, in exhibiting its own validity).

⁶Our less ambitious test sidesteps this discussion, since it is not intended as a perfectly general test of unformativeness, but only as a way of lassoing a particular manifestation of it. On the other hand it is compatible with ambitions towards greater generality, since maybe the test can be integrated into a clear and general account of the elusive informative/uninformative distinction.

out *this intention*. (Searle, 1983, p. 85)

If we suppose this to be a proposal for an account of acting (Searle does not explicitly say this),⁷ it is, first of all, clear that it describes an Anscombian situation: There's a kind of attitude (an intention on part of Searle) and a kind of doing (that Searle performs the action of raising his arm by way of carrying it out), where the former is necessary for the latter and the latter is the former's object. And my point about Searle is that his formulation of the view clearly allows an interlocutor to keep asking: What is the content of this intention to raise your arm? For the answer that its content involves "carrying out" this intention presumably means that what is done must accord with this intention's content — which is what, again? The interlocutor will not be helped by the idea, whatever it comes to, that *self-reference* is involved, or that, as Searle adds, this self-reference is "causal". These additions do nothing to solve the infinitely recurring question of what we're supposed to take as the content of this supposedly causally self-referential intention.

For another example, Setiya, concerned with specifying an attitude of "taking as a reason" he thinks essential to acting for a reason, says the following:

[T]he attitude of taking p as my reason to act must present *itself* as part of what motivates my action. The content of taking-as-one's-reason is thus *self-referential*: in acting [for the reason that] p , I take p to be a consideration belief in which motivates me to ϕ *because I so take it*. (Setiya, 2007, p. 45)⁸

One might think that Setiya's proposal does not place us in an Anscombian

⁷Searle can be taken as saying that he is taking both action and intention for granted, and wants only to account for the way in which they relate. Since the way in which they relate involves satisfying an intention, which involves acting, the charge of unformativeness below does, I think, affect even this ambition.

⁸I prefer this formulation, since Setiya's final refinement of the view replaces the "because I so take it" bit with "hereby", which seems to hide the self-referential aspect of the attitude from view.

situation since he does not explicitly say that this intricate attitude, required for acting for a reason, represents someone as *acting for a reason*. Actually, Setiya says in a footnote, perhaps out of some desire to avoid worries about circularity, that “in taking p as one’s reason for doing ϕ , one believes something whose truth entails that one is acting for that reason — not [...] the very proposition that one is doing ϕ for the reason that p ” (Setiya, 2007, p. 46, n. 39). I’ve said that I do not find such claims readily acceptable. For what could it be to believe something that *entails* that one acts for a reason, without *believing* that one acts for a reason? But we need not get bogged down in this matter, for it is clear that the self-reference involved in Setiya’s proposal generates an Anscombian situation in a closely related but subtly different way.

Setiya clearly does say that the attitude of *taking p as a reason* is satisfied in part by the fact that p is taken as a reason. In fact it seems to be a large part of the point of Setiya’s view to introduce an attitude of taking something — p — as one’s reason, and to say that this attitude represents one as being *motivated* by the fact *that one takes p as one’s reason*. This manner of speaking makes it clear that a case of taking-as-one’s-reason requires a representation of a case of taking-as-one’s-reason. Now it seems equally clear that our test of uninformativeness applies, since nothing in this proposal answers the question of just how a subject thinks of p when they are taking themselves (as taking themselves as taking themselves... and so on forever) to take p as the reason for their action. The worry is not alleviated by allowing, as is important for Setiya (2007, pp. 44-45) that taking oneself to take p as one’s reason involves taking p as motivational in a causal sense.⁹

⁹Curiously, Setiya raises a reminiscent point about circularity against an account given by Marcus (2012):

The problem with this account is its patent circularity. What it is to act because p is to know that one is acting because p : the explanatory relation figures in a statement of its own nature. No-one who accepts as substantive the need to explain necessary truths can tolerate circularity of this kind in accounts of what it is to ϕ . (Setiya, 2013, p. 511)

The above hopefully shows that, if a theorist introduces a situation that fits Anscombe's description, then this theorist can't respond to the charge of unformativeness simply by adding descriptions — even if they are necessary or essential — of the kind of thing that is the object of the attitude introduced. But all this will, of course, have us wondering how an account describing an Anscombian situation could possibly escape the charge of unformativeness: How could an account be circular in the way Anscombe describes and still not invite the same question again and again? What we've learned so far is this: In giving our account, we can't just keep mentioning the object of the attitude, since that reintroduces the attitude, which reintroduces its object, and so on forever, and we can't just keep characterising that object in a partial way which *doesn't* reintroduce this object, since that never fully answers the question. But what else could we say?

My idea is that, to ward off the appearance that there is some question our account leaves unanswered, we need an insight about its object which allows us to give a *new answer* whenever a question of the form "what is X?" is introduced. And I think that we have such an insight in a previously discovered thesis about the instrumental organisation of a thought about acting. For I earlier defended the following three-way equivalence:

Equivalence N believes N is acting in Aing = N believes N is Xing in order to A = N believes N is acting in (Xing in such a way that it causally contributes to Aing)

Our present concern about informativeness seems to affect Marcus and Setiya equally. In fact I am not sure why Setiya's worry about circularity shouldn't affect both accounts as well. For even if Setiya is not committed to saying that acting for the reason that p involves believing one is acting for the reason that p, he clearly seems to introduce the "explanatory connection" he thinks proper to action in his account of it, since to him that explanatory connection (involving, as it does, taking something as one's reason) involves thought about that explanatory connection. Perhaps what's supposed to save Setiya and not Marcus is that Setiya's account is not "immediate". But it is not clear to me what makes it less immediate, nor just what makes immediacy a virtue.

In this we've given an account of acting in Aing, among the conditions of which appears the statement that the subject must believe they are acting in Aing. Imagine that an interlocutor bent on charging us with un informativeness asks: "But what's that?". It is quite clear that the above thesis allows us now to generate a new answer, which is not simply a partial characterisation of the object of the belief. For may say: Believing you're acting in Aing is believing you're doing something else — Xing — *in order to A*. And now we see quite clearly a way in which this piece of dialogue could keep going. The interlocutor might now ask what's thought about in a belief about Xing in order to A. But here again we have an answer: "It's acting in Xing, you see", and if a further question of the same sort is pressed, a further answer of the same sort could be given. The thesis above allows the dialogue to go on in such a way that no question ever needs to remain unanswered. And if there is no question which need be left unanswered by the account, it is hard to see how it could be right to say that this account fails to say what its subject matter is.

Bibliography

- Alvarez, M. (2009). Acting Intentionally and Acting for a Reason. *Inquiry*, 52(3):293–305. 129
- Alvarez, M. (2010). *Kinds of Reasons*. Oxford University Press. 129
- Alvarez, M. and Hyman, J. (1998). Agents and their actions. *Philosophy*, 73(02):219–245. 54, 55
- Anscombe, G. E. M. (1963). *Intention*. Harvard University Press. 2, 3, 26, 38, 55, 68, 73, 74, 76, 77, 88, 89, 95, 96, 112, 115, 158
- Anscombe, G. E. M. (1969). On Promising and Its Justice, and Whether It Needs be Respected In Foro Interno. *Crítica: Revista Hispanoamericana de Filosofía*, 3(7/8):61–83. 177
- Bishop, J. (1987). Sensitive and insensitive responses to deviant action. *Australasian Journal of Philosophy*, 65(4):452–469. 50
- Bishop, J. (2012). Exercising Control in Practical Reasoning: Problems for Naturalism About Agency. *Philosophical Issues*, 22(1):53–72. 21, 30
- Bishop, J. C. (1989). *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge University Press. 48

- Bishop, J. C. (2011). Review of Life and Action: Elementary Structures of Practice and Practical Thought. *Ethics*, 122(1):212–220. 96
- Braun, D. (2015). Desiring, Desires, and Desire Ascriptions. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 172(1):141–162. 138, 144
- Brewer, T. (2006). Three Dogmas of Desire. In *Values and Virtues: Aristotelianism in Contemporary Ethics*. Oxford University Press. 140
- Castañeda, H.-N. (1979). Intensionality and Identity in Human Action and Philosophical Method. *Noûs*, 13(2):235–260. 97
- Castañeda, H.-N. (1992). Indexical Reference and Bodily Causal Diagrams in Intentional Action. *Studia Logica: An International Journal for Symbolic Logic*, 51(3/4):439–462. 97
- Choi, S. (2008). Dispositional Properties and Counterfactual Conditionals. *Mind*, 117(468):795–841. 61
- Churchland, P. M. (2014). Consciousness and the Introspection of ‘Qualitative Simples’. In Brown, R., editor, *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*, number 6 in *Studies in Brain and Mind*, pages 35–56. Springer Netherlands. 114
- Coope, U. (2007). Aristotle on Action. *Aristotelian Society Supplementary Volume*, 81(1):109–138. 2, 20, 65
- Crane, T. (2001). *Elements of mind*. Oxford University Press Oxford. 140
- Dancy, J. (2000). *Practical Reality*. Oxford University Press. 39, 128, 129
- Dancy, J. (2004). Two Ways of Explaining Actions. *Royal Institute of Philosophy Supplements*, 55:25–42. 130

- Danto, A. C. (1965). Basic Actions. *American Philosophical Quarterly*, 2(2):141–148. 104, 105
- Danto, A. C. (1981). *The Transfiguration of the Commonplace: A Philosophy of Art*. Harvard University Press. 25
- Darwall, S. (2003). Desires, Reasons, and Causes. *Philosophy and Phenomenological Research*, 67(2):436–443. 129
- Davidson, D. (2001a). *Essays on Actions and Events: Philosophical Essays*. Oxford University Press. 9, 18, 27, 28, 32, 37, 39, 41, 43, 78, 80, 83, 130, 131, 132
- Davidson, D. (2001b). *Inquiries into truth and interpretation*. Clarendon Press. 178
- Davidson, D. (2004). *Problems of Rationality*. Clarendon Press. 2, 43, 122
- Davis, W. A. (2005). Reasons and psychological causes. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 122(1):51–101. 39, 129
- Davis, W. A. (2010). The Causal Theory of Action. In O'Connor, T. and Sandis, C., editors, *A Companion to the Philosophy of Action*, pages 32–39. Wiley-Blackwell. 21
- Devitt, M. (1997). *Realism and Truth*. Princeton University Press. 175
- Dummett, M. (1991). *The logical basis of metaphysics*. Duckworth, London. 179
- Falvey, K. (2000). Knowledge in Intention. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 99(1):21–44. 11, 76

- Fara, D. G. (2003). Desires, Scope, and Tense. *Philosophical Perspectives*, 17(1):141–163. 142
- Fara, D. G. (2013). Specifying Desires. *Noûs*, 47(2):250–272. 137, 144, 145, 146, 148, 150
- Fara, M. (2005). Dispositions and habituals. *Noûs*, 39(1):43–82. 61, 62
- Ford, A. (2011). Action and Generality. In Ford, A., Hornsby, J., and Stoutland, F., editors, *Essays on Anscombe's Intention*, pages 76–104. Harvard University Press. 24, 176
- Ford, A. (2015). The Arithmetic of Intention. *American Philosophical Quarterly*, 52(2):129–143. 115
- Frankfurt, H. G. (1978). The Problem of Action. *American Philosophical Quarterly*, 15(2):157–162. 53, 60
- Ginet, C. (1975). *Knowledge, Perception and Memory*. Springer Science & Business Media. 112
- Ginet, C. (1990). *On Action*. Cambridge University Press. 167
- Goldman, A. I. (1970). *A theory of human action*. Prentice-Hall. 43
- Haack, S. (1987). Realism. *Synthese*, 73(2):275–299. 175
- Haddock, A. (2011). The Knowledge That a Man has of His Intentional Actions. In Ford, A., Hornsby, J., and Stoutland, F., editors, *Essays on Anscombe's Intention*, pages 147–169. Harvard University Press. 74
- Haddock, A. (2012). Meaning, Justification, and 'Primitive Normativity'. *Aristotelian Society Supplementary Volume*, 86(1):147–174. 176
- Harman, G. (1976). Practical Reasoning. *The Review of Metaphysics*, 29(3):431–463. 6, 167, 169, 180

- Hieronymi, P. (2009). Believing at Will. *Canadian Journal of Philosophy, Supplementary Volume*, 35:149–187. 17
- Hornsby, J. (1980). *Actions*. Number 124. Routledge and Kegan Paul. 104, 110
- Hornsby, J. (2004). Agency and actions. *Royal Institute of Philosophy Supplements*, 55:1–23. 20, 23, 26
- Hornsby, J. (2008). Agency and alienation. In de Caro, M. and MacArthur, D., editors, *Naturalism In Question*, pages 173–187. Harvard University Press. 2, 20
- Hornsby, J. (2012). Actions and activity. *Philosophical Issues*, 22(1):233–245. 10, 11
- Hornsby, J. and Stanley, J. (2005). Semantic Knowledge and Practical Knowledge. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 79:107–145. 110
- Hyman, J. (2014). Desires, Dispositions and Deviant Causal Chains. *Philosophy*, 89(01):83–112. 2, 56, 65
- Kapitan, T. (1995). Intentions and Self-Referential Content. *Philosophical Papers*, 24(3):151–166. 173
- Keefe, R. (2002). XV—When does Circularity Matter? *Proceedings of the Aristotelian Society*, 102(1):275–292. 180
- Lavin, D. (2013). Must There Be Basic Action? *Noûs*, 47(2):273–301. 106, 107
- Levin, J. (2013). Functionalism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2013 edition. 51
- Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50(3):249–258. 51

- Lewis, D. (1997). Finkish dispositions. *The Philosophical Quarterly*, 47(187):143–158. 58
- Lewis, D. K. (1966). An argument for the identity theory. *The Journal of Philosophy*, 63(1):17–25. 51
- Lycan, W. G. (2012). Desire Considered as a Propositional Attitude. *Philosophical Perspectives*, 26(1):201–215. 137, 154, 155
- Marcus, E. (2012). *Rational Causation*. Harvard University Press. 182
- Martin, C. B. (1994). Dispositions and Conditionals. *The Philosophical Quarterly*, 44(174):1–8. 61
- McDaniel, K. and Bradley, B. (2008). Desires. *Mind*, 117(466):267–302. 153
- McDowell, J. (1996). *Mind and World*. Harvard University Press. 178
- McDowell, J. (2010). What Is the Content of an Intention in Action? *Ratio*, 23(4):415–432. 74
- Mele, A. (1999). Is there a place for intention in an analysis of intentional action? *Philosophia*, 27(3-4):419–432. 167
- Mele, A. R. (1987). Are Intentions Self-Referential? *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 52(3):309–329. 167, 168
- Mele, A. R. (1992). *Springs of Action: Understanding Intentional Behavior*. Oxford University Press, New York. 47, 57
- Merricks, T. (2009). Propositional Attitudes? *Proceedings of the Aristotelian Society (Hardback)*, 109:207–232. 140
- Millar, A. (2004). *Understanding People: Normativity and Rationalizing Explanation*. Clarendon Press. 129

- Montague, M. (2007). Against Propositionalism. *Noûs*, 41(3):503–518. 140
- Mourelatos, A. P. D. (1978). Events, processes, and states. *Linguistics and Philosophy*, 2(3):415–434. 11, 14, 15, 16
- Nagel, T. (1970). *The Possibility of Altruism*. Oxford Clarendon Press. 123, 124, 125, 126, 127
- Nagel, T. (1986). *The View From Nowhere*. Oxford University Press. 20, 21
- Paul, S. K. (2009a). How we know what we’re doing. *Philosophers’ Imprint*, 9(11):1–24. 74
- Paul, S. K. (2009b). Intention, Belief, and Wishful Thinking: Setiya on “Practical Knowledge”. *Ethics*, 119(3):546–557. 74, 81
- Paul, S. K. (2011). Deviant Formal Causation. *Journal of Ethics and Social Philosophy*, 5(3):1–23. 96, 97, 98
- Peacocke, C. (1979). *Holistic Explanation: Action, Space, Interpretation*. Clarendon Press. 48, 49, 50, 51, 71
- Quine, W. V. O. (1960). *Word and Object*. Technology Press of the Massachusetts Inst. of Technology. 178
- Rödl, S. (2007). *Self-Consciousness*. Harvard University Press. 2, 11, 12
- Ryle, G. (1949). *Concept of Mind*. Penguin. 13, 60, 120, 121, 173
- Sartorio, C. (2009). Omissions and causalism. *Noûs*, 43(3):513–530. 27
- Searle, J. R. (1980). The Intentionality of Intention and Action*. *Cognitive Science*, 4(1):47–70. 25
- Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press. 6, 140, 167, 168, 180, 181

- Sehon, S. R. (1997). Deviant causal chains and the irreducibility of teleological explanation. *Pacific Philosophical Quarterly*, 78(2):195–213. 47, 48, 50, 52
- Sellars, W. S. (1956). Empiricism and the philosophy of mind. *Minnesota Studies in the Philosophy of Science*, 1:253–329. 178
- Setiya, K. (2003). Explaining action. *The Philosophical Review*, 112(3):339–393. 42, 45
- Setiya, K. (2007). *Reasons Without Rationalism*. Princeton University Press. 6, 9, 18, 36, 38, 43, 46, 72, 129, 180, 181, 182
- Setiya, K. (2008a). Believing at will. *Midwest Studies In Philosophy*, 32(1):36–52. 17
- Setiya, K. (2008b). Practical Knowledge. *Ethics*, 118(3):388–409. 74, 78, 81, 87
- Setiya, K. (2009). Practical Knowledge Revisited. *Ethics*, 120(1):128–137. 74, 78, 81
- Setiya, K. (2011a). Knowledge of Intention. In Ford, A., Hornsby, J., and Stoutland, F., editors, *Essays on Anscombe's Intention*, pages 170–197. Harvard University Press. 74, 76
- Setiya, K. (2011b). Reasons and causes. *European Journal of Philosophy*, 19(1):129–157. 37, 39, 53, 54, 57
- Setiya, K. (2013). Causality in Action. *Analysis*, 73(3):501–512. 182
- Sinhababu, N. (2009). The humean theory of motivation reformulated and defended. *Philosophical Review*, 118(4):465–500. 51
- Smith, M. (1987). The humean theory of motivation. *Mind*, 96(381):36–61. 36
- Smith, M. (1994). *The Moral Problem*. Blackwell. 2

- Smith, M. (2009). The Explanatory Role of Being Rational. In Sobel, D. and Wall, S., editors, *Reasons for Action*, pages 58–80. Cambridge University Press. 117
- Smith, M. (2012). Four objections to the standard story of action (and four replies). *Philosophical Issues*, 22(1):387–401. 9, 25, 27, 48, 129
- Sober, E. (1982). Realism and Independence. *Noûs*, 16(3):369–385. 175
- Stanley, J. and Williamson, T. (2001). Knowing How. *The Journal of Philosophy*, 98(8):411–444. 110, 112, 173
- Stein, N. (2014). Causes and Categories. *Noûs*. 32, 37
- Stout, R. (1996). *Things That Happen Because They Should: A Teleological Approach to Action*. Oxford University Press. 91
- Stoutland, F. (1980). Oblique causation and reasons for action. *Synthese*, 43(3):351–367. 117
- Strawson, P. F. (1992). *Analysis and Metaphysics: An Introduction to Philosophy*. Oxford University Press. 179
- Thagard, P. (2006). Desires Are Not Propositional Attitudes. *Dialogue: Canadian Philosophical Review/Revue canadienne de philosophie*, 45(01):151–156. 140
- Thalberg, I. (1984). Do our intentions cause our intentional actions? *American Philosophical Quarterly*, 21(3):249–260. 57
- Thompson, M. (2008). *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press. 11, 54, 76, 78, 99, 108, 110, 140, 158, 176
- Velleman, J. D. (1989). *Practical Reflection*. Princeton University Press. 180

- Velleman, J. D. (1992). What Happens When Someone Acts? *Mind*, 101(403):461–481. 2, 18, 20, 35, 58
- Vendler, Z. (1957). Verbs and Times. *The Philosophical Review*, 66(2):143–160. 16
- Wilson, G. M. (1989). *The Intentionality of Human Action*. Stanford University Press. 52, 70, 72, 167
- Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus: German and English*. Routledge. 10
- Wittgenstein, L. (1953). *Philosophical Investigations*. B. Blackwell. 112, 119
- Wolfson, B. (2012). Agential Knowledge, Action and Process. *Theoria*, 78(4):326–357. 93