

Copyright
by
Jonathan Paul Drake
2017

The Dissertation Committee for Jonathan P. Drake Certifies that this is the approved version of the following dissertation:

THE (IR)RELEVANCE OF TRUTH TO RATIONALITY

Committee:

Jonathan Dancy, Supervisor

David Sosa, Co-Supervisor

Maria Alvarez

Daniel Bonevac

L. Ray Buchanan

THE (IR)RELEVANCE OF TRUTH TO RATIONALITY

by

Jonathan Paul Drake

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2017

Preface

I was first introduced to philosophical problems surrounding reasons just over five years ago. Jonathan Dancy held a seminar on action theory, which I attended in my first year of graduate school. As many first-year graduate students do, I was struggling a bit to keep up. But some time well into the course, he introduced a debate over a particular question — the answer to which I thought was perfectly obvious. He put the question to us like this. Imagine that it is a fine winter day, and you live near a pond. Occasionally you like to sit by the frozen pond and watch people skate on it. Today is such a day, and there is only one person skating: Edmund. After a few minutes, you notice that Edmund skates only near the edge of the pond. You approach him and ask him, “Why do you keep to the edge of the pond?” He responds by telling you that the ice in the middle of the pond is thin. You find out later, though, that the ice in the middle of the pond is not thin.

In normal cases, when someone acts for the reason that (for example) the ice in the middle of the pond is thin, it really is the case that the ice in the middle is thin. This is mostly due to the fact that we are not often wrong about such mundane ways the world is. But, Jonathan asked, what if one takes it that the ice is thin, and in fact it is not thin? Can one still act *for the reason* that the ice is thin? I immediately thought that the answer is very clearly, “Yes!” “Just yesterday,” I remember thinking to myself, “my wife went to the store to get a particular ice cream that she likes, only to discover that they did not have it.” Surely she still went there to get that ice cream, or alternatively, for the reason that they have the ice cream that she likes. I expect that many of you are like I was in finding nothing suspicious about saying such things.

I was surprised to hear that Jonathan was all but alone in defending this “Nonfactive View” of acting for a reason. At that juncture, I could not even understand why one would be tempted to hold the opposing view. But it was not long before I was confronted with a handful of puzzles that emerges when you answer these questions in these ways (and I was given a long working over by the likes of Jonathan Dancy, David Sosa, Ray Buchanan, Dan Bonevac, Maria Alvarez, Tim Williamson, Clayton Littlejohn, Michael Bratman, Stephen Darwall, Juan Comesaña, Miriam Schoenfield, and others). Chief among these puzzles, I thought, was the following. The Nonfactive View of acting for a reason says that an agent can perform an action for a reason even if that reason is something that is not the case. It is a very plausible thought that whenever someone acts for a reason, that reason can serve to explain their doing what they do. If you unexpectedly found me in the Department this week, you might have asked me why I have come here. It is very natural to take my answer to this ‘Why?’ question — whatever it may be — as both an explanation of my action and as a statement of the reason for which I acted. I might have said, for example, that “I defend my dissertation on Friday morning.” Here comes the trouble, though. A second very plausible thought is that anything that explains something must be the case; something that is not the case cannot explain anything. A flying baseball cannot explain a broken window if there is, in fact, no flying baseball. Perhaps you can already see the trouble brewing. If reasons can explain, and things that explain are always the case, then it seems to follow straightaway that reasons must be things that are the case (and so the Nonfactive View cannot be true).

When Jonathan confronted me with this puzzle, I was immediately enthralled by it and its surrounding issues. Very quickly I found that thinking about reasons took me to the heart of foundational issues in ethics, epistemology, and metaphysics. At first, I began to see far too many of these foundational issues as nails, and my solution to the puzzle as a

hammer. But I was gradually disabused of this grand notion — mostly thanks to Maria Alvarez. I came to see that my real interest was the rational evaluation and explanation of agents and their actions (broadly construed, so as to include epistemic actions such as forming a belief), and that I was interested in a fairly specific set of questions in this area. But I also came to think that the notion of a reason had the potential to get philosophers working in different areas speaking the same language; and I developed a hope (which I still hold onto) that illumination of this notion would put philosophers in these different areas in a position to more readily exchange insights.

In the end, I have come to think that my initial reaction to Jonathan’s puzzle was mostly correct. But I was quite wrong to think that there was no reason to hold otherwise, and I have come to appreciate the good cases for a number of sophisticated views in the area. I set myself to defend a generally “nonfactive” view of reasons from a handful of forceful objections having to do with explanation and rationality. My efforts have led me to a package of views that spans several areas of inquiry, and one which I think is genuinely novel. If there is one theme to this package of views, it is that truth is irrelevant to rationality and justification in significant and surprising ways.

In completing this project I have become indebted to a great many people. The staff in the graduate program at the University of Texas at Austin have been invaluable to me — especially Michelle Botello, Sally Jackman, and Stephanie Hollub-Fletcher. I am appreciative of audiences at the Factive Turn Workshop in Vienna, the 2016 Pacific Division Meeting of the APA, the Practical Reason and Metaethics Workshop at the University of Nebraska, and several forums at the University of Texas at Austin. I have benefited from helpful conversation with Michael Bratman, Brian Cutter, Stephen Darwall, Sinan Dogramaci, Jeremy Evans, Jon Litland, Clayton Littlejohn, Brian Miller, Miriam Schoenfield, Jason Schukraft, and Tim Williamson. I am thankful for feedback

on drafts from Juan Comesaña, Veli Mitova, and various anonymous referees. I am grateful to each of my committee members, who have done all of this and more: Dan Bonevac and Ray Buchanan helped me to see the importance of purposes and infinitivals to my project; and David Sosa and Maria Alvarez helped me to genuinely appreciate and respect my opposition. I am especially grateful to Jonathan Dancy, who, more than anyone, has taught me how to think about these issues. I owe a unique debt to Daniel Eaton, who has helped me in ways that I cannot adequately express. Finally, I would not have got along at all without the unfailing support of my family — especially my wife, Emily.

The (Ir)Relevance of Truth to Rationality

Jonathan Paul Drake, Ph.D.

The University of Texas at Austin, 2017

Supervisor: Jonathan Dancy

Co-Supervisor: David Sosa

It is possible to act for a reason. We do it all the time. You might have brought her medicine for the reason that she is ill. He might go to the store to get milk. Edmund might skate in the middle of the pond because the ice in the middle of the pond is thin. What must be true of us, and of the world, such that we can act for reasons?

In normal cases, when someone acts for the reason that (for example) the ice in the middle of the pond is thin, it really is the case that the ice in the middle is thin. This is mostly due to the fact that we are not often wrong about such mundane ways the world is. But what if one takes it that the ice is thin, and in fact it is not thin? Can one still act *for the reason* that the ice is thin?

In my efforts to give a sufficient answer to this question, I have been led to a package of views, the core tenets of which are at least the following five. First, it is possible to act in the light of a falsehood: a consideration that is not the case can be an agent's reason for acting. Second, it is not possible to act in unbelief: in order for an agent to act for a reason, the agent must at least believe that reason to be the case. Third, the reasons for which agents act can play a role in explaining the actions done for those reasons — even when agents act in the light of falsehoods. Fourth, there are very few (if any) formal rules or principles constraining the explanatory role of reasons. Any action explanation

role for the reason. Fifth, all of these claims apply equally to motivating and normative reasons, so-called practical and epistemic reasons, and reasons for action and reasons for belief.

Table of Contents

Chapter One: Acting in the Light of a Falsehood	1
Acting For a Reason: The Very Idea.....	4
Reasons, Explanations, and Factivity	22
The Argument from the Factivity of Explanation	26
Contributing to an Explanation.....	52
Factual Explanations for the Nonfactive View: A Brief Sketch	57
Chapter Two: Acting in Unbelief	63
Clearing Ground for the Question of the Cognitive Constraint.....	65
An Initial Defense of the Belief View.....	79
Acting in Ignorance	92
Acting in Unbelief.....	109
Are Reasons Belief?.....	114
Chapter Three: Arguments from Error	138
The Use of Error Cases in Action Theory.....	140
Arguments from Factual Error	146
Arguments from Procedural Error.....	159
Variations in Explanation	173
Summary Remarks	194
Chapter Four: Motivating Reason to Slow the Factive Turn in Epistemology..	203
Preliminary Remarks	205
Uniformity.....	209
Acting Rationally in the Light of a Falsehood	212
Motivating Reason to Slow the Factive Turn.....	219
Reasons and Rationality	220
Can the Factive Turn Be Forced?.....	228
A Practical Theory of Epistemic Reasons: A Brief Sketch	236
References.....	240

Chapter One: Acting in the Light of a Falsehood

Edmund is about to go skating on a pond near his home. He takes it that the ice in the middle of the pond is dangerously thin. So, when he skates on the pond, Edmund keeps to the edge of the pond. You are on a nearby hill, observing the skaters on the pond as you enjoy the nice winter day. As Edmund finishes up skating, you approach him and ask him why he is skating only near the edge of the pond. He responds by telling you that the ice in the middle of the pond is thin. After this brief exchange with Edmund, you go out on the pond to skate. Initially, you keep to the edge; but then you gradually try your luck to see how the ice is in the middle of the pond. You find that the ice in the middle of the pond is not thin.

Here are a few intuitive claims about this story. First, when Edmund answers your ‘Why?’ question concerning his skating, he is articulating to you his reason for doing what he was doing. Second, in the case, you would readily accept Edmund’s answer; and you would reasonably believe, on the basis of what he says to you, that his reason for skating near the edge of the pond is that the ice in the middle of the pond is thin. Third, both you and Edmund would be accurate in speech and belief: Edmund’s reason for skating near the edge of the pond *really is* that the ice in the middle of the pond is thin. Finally, the last detail of the story does nothing to make any of these intuitive claims less compelling. Most importantly: the fact that Edmund is wrong about how things are with the ice does not make you wonder, as it were, what his reason *really* was.

If all of this is correct, it seems plausible to say that Edmund can act for the reason that the ice in the middle of the pond is thin even when it is not the case that the ice in the middle of the pond is thin. Generalizing, we could say that acting for a reason is not, as philosophers of action say, *factive*: it does not follow from the fact that Edmund acts for the

reason that r that it is the case that r . This is what the Nonfactive View of acting for a reason holds. So we could also say that the Nonfactive View seems to be the intuitive or default view concerning the factivity of acting for a reason.

It is perhaps surprising, then, that the Nonfactive View is by far in the minority among philosophers. Instead, the Factive View — which holds that the reasons for which agents act must be facts — is the dominant view. How could this be? The story we just told about Edmund seemed perfectly natural. I offer three quite general diagnoses. First, many have been influenced by the tradition which holds that the reasons for which agents act are psychological states of the acting agents, or some other feature of the world which causes the action. Since it is very plausible that only something that is the case can cause anything, it might thereby seem undeniable that reasons (as a subset of the things that are causes) must be things that are the case. Second, there is growing momentum behind the idea that there is a connection between knowledge and acting for a reason. Specifically, some have begun to advance the view that acting for a reason requires knowledge that that reason is the case. Since one cannot know anything that is not true, this way of thinking has led some to suppose that the reasons for which agents act must also be truths. Third, there seems to be an independently powerful argument against the intuitive picture just painted. The argument is based on a thought about the nature of explanation. It is a common thought that explanation is factive: only something that is the case can explain something else. It is also tempting to suppose that when an agent acts for a reason, that reason can serve to explain the action done for that reason. It may appear to follow, then, that when an agent acts for a reason, that reason must be something that is the case.

Of these three challenges to the intuitive picture of Edmund we were just considering, I find the third challenge to be the most compelling. This is partly because,

as we will see in the third chapter, the first two challenges tend to rely on somewhat robust theories including substantive assumptions about other subjects. The third challenge, by contrast, seems to stand on its own, adducing only considerations to which it would seem all parties would agree. Indeed, the ease with which the premises of the argument would be granted almost seems to suggest that the Nonfactive View of acting for reasons is a nonstarter. I believe that this dialectical state of affairs presents us with a puzzle about the nature of reasons, action, and explanation. Intuitively, as we just saw, one can act for the reason that r even when it is not the case that r . But the argument under consideration would seem to show that this intuitive claim is at odds with the conjunction of two *other* intuitive claims: that reasons explain actions, and that explanation is factive. In this first chapter I aim to dissolve this puzzle. While many theorists seem to think that the above argument shows that the reasons for which agents act must be facts, I shall argue that it fails to do so. What I believe has gone unnoticed about this argument is that it suffers from a fatal ambiguity in its central notion: the notion of something *explaining* something else. I shall suggest that, at the very least, we need a distinction between something being an explanation and something contributing to an explanation. Once this central notion is sufficiently disambiguated, though, the argument in question loses its apparent force.

The strategy that I am presenting here is quite different from the few already present in the literature. So far, those who have addressed this challenge on behalf of the Nonfactive View have been involved in a kind of damage control: in one way or another, they have suggested that things are not as bad off for the Nonfactive View as this argument might initially lead one to believe. Contrastingly, I think that there simply is no damage to control; this is simply a poor argument against the Nonfactive View, and there is no good argument of its kind forthcoming.

The chapter proceeds as follows. In the first section, I will lay out some preliminary notions which will stay with us for the remainder of this manuscript. I will also address the suggestion that the very idea of acting for a reason is factive. I consider the going views as to what a reason is, and I show that none of them are in themselves suggestive of a factive notion of acting for a reason (on the contrary, some are suggestive of a nonfactive notion). Nearby are considerations as to whether natural language constructions of the form “*A*’s reason was that *r*” (for example) can be true only if it is true that *r*. I argue that such constructions are also not factive. So, at best, conceptual and semantic analysis of the notion of acting for a reason is suggestive of the Nonfactive View; at worst, it is not suggestive of any view on the issue. Either way, such analyses leave the Nonfactive View unscathed. Enter the argument motivated by considerations of the nature of explanation. In the second section, I lay out some preliminaries that will allow us to understand this argument. In the third section, I demonstrate that the argument suffers from an ambiguity in the notion of *being explanatory*. I disambiguate five senses of the notion, and show that on each of these disambiguations, the argument that acting for a reason is factive loses its force. I conclude by sketching out some of the explanatory resources and schemata all of this leaves available to the Nonfactive View. I return to this issue in the third chapter, where I lay out a more complete picture of what action explanation can look like on a Nonfactive View of acting for a reason.

ACTING FOR A REASON: THE VERY IDEA

In asking if acting for a reason is factive — or whether the reasons for which agents act must be facts — it seems sensible to devote some time to the questions of what it is to act for a reason, and what a reason for acting is.

Kinds of Reasons

We can begin with a perfectly ordinary, and now standard, distinction between two kinds of reasons.¹ We often say that people do things *for* reasons: “She is doing it for the reason that it will make her son happy;” “His reason for doing it was to pay off his debt;” “They did it because he needed help.” When we talk (and think) this way, we invoke the notion of what philosophers now standardly refer to as a **motivating reason**. We will approach conceptual analyses of motivating reasons shortly; but we can say roughly that a motivating reason is a consideration that motivates an agent to act. Such reasons can often be summoned to explain actions done for reasons. For example, suppose you see me running through the halls. You might ask *why* I am running through the halls. I would likely respond by presenting the consideration(s) which motivated me to run through the halls (for example, the consideration that I am late to give me eleven o’clock lecture). You would likely be satisfied with such a response, as it would help you to make sense of my action. As we will see, there are legitimate doubts that motivating reasons can *always* serve to explain the actions they motivate. If they cannot, we may need to distinguish between motivating reasons and **explanatory reasons**: considerations that explain actions.

Even if motivating reasons and explanatory reasons are ontologically separable, they both seem capable of answering some sense of the question “Why is she doing that?” But we might be interested not only in why someone *is* doing what they are doing; we might also be interested in why, as it were, they *should* be doing what they are doing. In

¹Too many to name employ this common distinction between motivating and normative reasons, which is perhaps owed to Michael Smith [1992: 329]. For relevant introductions to the distinction, though, see Jonathan Dancy [2000: 1ff, 20-25] for something brief, and Maria Alvarez [2010; 2016] for something more thorough. Cases like the ones central to this project might impress upon us the idea that we do ultimately need a third class: explanatory reasons. Alvarez [2010; 2016] has done well to argue for this position, which I will address along the way.

addition to wondering what my reason is for running through the halls, you might also be interested in whether there is, as we might say, any *good* reason for me to be running through the halls. Here we are interested in reasons for acting in yet another sense. We often say that people *do things for* reasons; but we also say that *there are reasons* for doing things: “One reason to do it now is that the bank will be closed tomorrow;” “That was no reason to pop his balloons!;” “You should do it this way because that will make for a beautiful painting.” When we talk this way, we invoke the notion of what philosophers refer to as a **normative reason**: a consideration that counts in favor of a response. In normal cases (such as the one considered here), the agent’s motivating reason will be a normative reason. In telling you that I am late to give my lecture, I have presented you with that consideration which cast my action in a favorable light. I have told you not only *why I am* running, but also, perhaps, *why I should be* running.

Recently, there has been a debate concerning the “factivity” of each of these kinds of reasons. In epistemology, there has been much more debate about the factivity of normative reasons. In that debate, there are two main views. One is the

Factive View_N: in order for r to be a reason for A to Φ , it must be the case that r .

The competitor to this view is the **Nonfactive View_N**, which is just the denial of the Factive View_N. In this context, the phrase “be a reason for A to Φ ” should not be taken not motivationally, but rather as, roughly, “count in favor of A ’s Φ ing.” According to the Factive View_N, my being late cannot count in favor of my running through the halls unless I really am late. Those who endorse the Nonfactive View_N will deny this claim, holding that my being late might count in favor of my running regardless of whether I really am late.

I will put aside for now the debate about the factivity of normative reasons. We will return to this debate in the fourth chapter. Our present concerns are about the factivity of motivating reasons, which has received much more attention in practical philosophy and action theory. In this debate, there are also two main views. One is the

Factive View_M: in order for A to Φ for the reason that r , it must be the case that r .

Adherents to this view include Maria Alvarez, Donald Davidson and his followers, Jennifer Hornsby, John Hyman, Clayton Littlejohn, John McDowell, Peter Unger, and Timothy Williamson and his followers.² The competitor to this view is the **Nonfactive View_M**, which is just the denial of the Factive View_M. The Nonfactive View_M is defended most prominently by Jonathan Dancy; but it is perhaps also defended by Juan Comesaña and Matthew McGrath, David Enoch, Mark Schroeder, and Kieran Setiya.³ In the context of this debate, “ Φ for the reason that r ” should not be taken normatively (as in “ r is a consideration that favors Φ ing”), but rather motivationally. So, independent of what (normative) reasons there are for me to be running through the hall, proponents of the Factive View_M will hold that I cannot be running through the halls *for the reason that* I am late unless I really am late. The Nonfactive View_M denies this, and holds that my reason

² See Alvarez [2009; 2010], Davidson [1963; 1980], Hornsby [2007; 2008], Hyman [1999; 2011; 2015], Littlejohn [2012], McDowell [2013], Unger [1975], Williamson [201X], and also Ian Schnee [2015].

³ See Dancy [2000; 2003; 2008; 2011; 2014], Comesaña and McGrath [2014], Enoch [2010] Schroeder [2008a], Setiya [2007], and perhaps Constantine Sandis [2013]. Comesaña, McGrath, and Schroeder all defend a nonfactive view of *having* reasons; the connection to acting for a reason is not completely clear. Per Alvarez [2016: 25], the label may be extended to those who, like Hornsby and McDowell, seem to hold what has been called Disjunctivism about acting for a reason. For reasons that will be made clear in the second and third chapters, I do not think it appropriate to lump Disjunctivists in with Nonfactivists.

for running might be that I am late even if I am not really late. This debate — between the Factive View_M and the Nonfactive View_M — is the one I turn to now.

What Is A Motivating Reason?

With these preliminaries on hand, we can now proceed with the questions of what it is to act for a reason, and what a (motivating) reason for acting is. It may seem simpler to start with the latter notion, since the former notion appears to contain the latter notion as a part: before we can say what it is to act for a *reason*, we should like to know just what a reason is. For purposes of this project, we should also like to be on the lookout for whether any analysis of these notions is suggestive of either the Factive View_M or the Nonfactive View_M .

Motivating reasons are often discussed as the *reasons for which* agents act. But analysis of this notion does not bear much fruit, not least because it itself makes use of the notion of a reason. I think it more helpful to start with the thought that motivating reasons are *considerations* which motivated the agent to act. In these contexts, the word “consideration,” like the word “belief,” admits of disambiguation. To be precise, we might say that motivating reasons are the things which, when considered, motivate the agent to act. Now, one standard way of characterizing the reasons for which agents act is as *considerations in the light of which agents act*.⁴ This metaphorical analysis, too, may not prove very useful in asking whether motivating reasons are facts. For to do something *in the light of* something else does not seem to amount to much more than doing something, *taking that something else into consideration*. In discussing motivating reasons — things of the sort involved in the earlier examples of skating around ponds and running through halls — we are surely looking for something more than that which is simply *taken into consideration*

⁴ See Alvarez [2010: 7, 35; 2016: 3] and Dancy [2000: 1].

when an agent acts. We are looking for something that has some motivational force.

Dancy elaborates on the notion this way:

... normally there will be, for each action, the reasons in the light of which the agent did that action, which we can think of as what persuaded him to do it.⁵

Under this characterization, motivating reasons are the things which, when considered by the agent, *persuade* the agent to do what she does. The language of persuasion is suggestive here: we are not looking for things that move agents to act in the sense that they merely *cause* them to act. We are looking, rather, for things that move agents to act in the rational sense. Hence even Davidson, who settled on the view that motivating reasons were, after all, causes, acknowledges:

What is the relation between a reason and an action when the reason explains the action by giving the agent's reason for doing what he did? We may call such explanations *rationalizations*, and say that the reason *rationalizes* the action... A reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action.⁶

Even if motivating reasons turn out to be causes, it seems clear that what we are interested in the first instance are not *mere* causes, but instead something like Davidson's *rationalizing* causes.

So we have as one analysis of the notion of a motivating reason: for r to be a reason for which $A \Phi$ s is for r to be a consideration that persuaded (or contributed to persuading) A to Φ . There are two other standard analyses of the notion that I should like to discuss. The first of these is the notion of a *consideration taken by an agent to favor an action*. This may seem very near to the previous analysis. After all, it might be the case that an agent's being persuaded to act by a consideration at least depends on the agent's taking

⁵ See Dancy [2000: 1].

⁶ See Davidson [1963: 685].

that consideration to favor the action in question. Even if that is so, this analysis is worth distinguishing, if only because it makes use of a normative concept — that of *favoring*. On this analysis, if $A \Phi$ s for the reason that r , r is a consideration that A took to be a (normative) reason for Φ ing.⁷ Motivating reasons are not simply the considerations that motivated the agent to act; they are the considerations that reveal the good the agent saw in doing what she did.

A third way of conceiving of motivating reasons is as considerations adduced as premises in practical reasoning.⁸ Jane, for example, might be thirsty, and so go to the fridge to get a drink. On this conception, the motivating reasons are the considerations Jane made use of in deciding to go to the fridge. She may have reasoned: “I am thirsty, and I should get a drink to quench my thirst. There are drinks in the fridge. So, I will go to the fridge.” In this (admittedly over-intellectualized) bit of reasoning, the considerations that Jane is thirsty, that a drink would quench her thirst, and that there are drinks in the fridge might all count (separately or together) as reasons for which Jane went to the fridge.

With a few standard analyses of what a motivating reason is in hand, we can pause now to ask: do any of the standard analyses of a motivating reason suggest that motivating reasons are facts?⁹ I think it is clear that, quite the contrary, these analyses are suggestive of a nonfactive notion. The consideration that she is ill, for example, can

⁷ See Pamela Hieronymi [2011: 411] for careful work in carving out this distinction. For instances of this standard usage, see Alvarez [2016: 3], Dancy [2000: 1ff], Alan Gibbard [1990: 162], Derek Parfit [1997; 2001], and Schroeder [2008]. Perhaps see also T. M. Scanlon’s [1998: 19] “operative reasons” as well as Joseph Raz [1975: 33].

⁸ See Alvarez [2016: 3], and especially the exchange between Dancy [2011] and Hyman [2011] invoking this conception. Alvarez [2016: 3] adds the qualification: “a premise in practical reasoning, if any, that leads to the action.”

⁹ I put aside for the moment analyses of motivating reasons as *explanatory* reasons, and as *facts which guide the agent in acting*. The former I put aside only for a moment, since the chapter is just about in what sense motivating reasons are explanatory. The latter I put aside because there is no sense in asking whether the notion of a guiding *fact* is factive. The more significant question is whether this is genuinely an analysis of motivating reasons, or of something else. I address that question in the second and third chapters.

clearly be a thing taken into consideration even if it is not the case that she is ill. The idea of persuasion does nothing to change this. We could ask Jane, for example, what consideration persuaded her to go to the fridge. She might cite the consideration that there are drinks in the fridge. Now, if it turns out that there are no drinks in the fridge, this is clearly not in conflict with the idea that that was the consideration that persuaded Jane to go to the fridge. Nearby is the idea of a premise of practical reasoning. Again, there is no suggestion of factivity here. We saw that Jane could reason that “There are drinks in the fridge; so I will go to the fridge.” There not being drinks in the fridge clearly does not prevent Jane from adducing this consideration as a premise in her reasoning about whether to go to the fridge. Finally, the notion of a consideration taken to favor actions likewise contains no suggestion of factivity. What considerations did Jane take to favor her going to the fridge? That there are drinks in the fridge was surely among them – this consideration reveals (at least part of) the good Jane saw in going to the fridge. But this does not imply that there actually *are* drinks in the fridge.

Does any of this change when we begin to construct, making use of these analyses, standard conceptions of what it is to *act for a reason*? I think not. Now, a common analysis of acting for a reason is that of *acting in the light of a consideration*. Despite its commonality, I am doubtful that this notion reveals much about the nature of the phenomenon in question. As I implied a moment ago, it is not altogether clear what it amounts to to act *in the light* of something over and above *taking that thing into consideration* in acting. But, as we just saw, the concept of a motivating reason surely includes more than this. Instead, I would like to start with the notion of *responding to a consideration*. I think it will be uncontroversial — perhaps even trivial — to say that acting for a reason is a way of responding to a consideration. Now, it is obvious that one can respond to a thing considered even when that thing is not the case. Suppose that Aria, a friend of yours, has

called off her wedding. You tell me that she has called off the wedding, and I ask for what reason she has done so. You tell me that she has done it in response to the consideration that her fiancé has been unfaithful. There is clearly no implication here that her fiancé actually has been unfaithful. This is perhaps evidenced by the fact that it would not be at all confusing or inappropriate for me to follow up by asking whether her fiancé has in fact been unfaithful.

So, that A responded to the consideration that r clearly does not entail that r . Does it make any difference if we elaborate on the notion of *responding to a consideration* by substituting in one of our standard analyses as to what sort of consideration a motivating reason is? I think not. It makes no difference, for example, if we amend the earlier story by saying that Aria was responding to the consideration which persuaded her to call off the wedding, namely, the consideration that her fiancé has been unfaithful. We could also amend the story by saying that Aria was responding to the consideration which she took to favor calling off the wedding, namely, the consideration that her fiancé has been unfaithful. Still, there is no implication that her fiancé has in fact been unfaithful. It is the same for the claim that Aria was responding to the considerations she took as premises in her reasoning about whether to call off the wedding, one of which was the consideration that her fiancé has been unfaithful.

If any of these standard analyses is complete and correct, then, it is already difficult to see how the Factive View_M could be true. Of course, these theories may be incorrect. Perhaps there is more to acting for a reason than responding to a consideration. But it is worth keeping in mind that we have not left the notion completely bare. We have elaborated the notion so that it makes use of the standard views as to *what sort* of consideration a motivating reason is. Even these elaborations contained no hint of factivity. I think that, in the light of these considerations, it is fair to say that the

conceptual analysis of a motivating reason and of acting for a reason provides no evidence for the Factive View_M. Quite the contrary, they might provide evidence for the opposing Nonfactive View_M. According to Maria Alvarez, for example,¹⁰

...reasons must be capable of being premises, i.e. things we reason, or draw conclusions, from, whether in theoretical or practical reasoning. Otherwise, the connection between reasons and reasoning would be lost.

But, as John Hyman notes,

...if by a person's reason we meant the premise or assumption on which he acted.... We can reason from false premises... For example, one can reason from the premise that the train from Paris will arrive on platform eight, even if in fact it will arrive on platform one.

If motivating reasons *must* be capable of being premises, and premises are the kinds of things that can be falsehoods, it might seem to follow that reasons must be capable of being falsehoods. This is just what the Nonfactive View_M says. Until we have in hand a critique as to why these standard analyses are misguided, I think we should be compelled to take the Nonfactive View_M as the default view in our theorizing about the factivity of reasons.

Are Reasons Contexts Factive?

Evidence from conceptual analysis notwithstanding, there are some who claim that certain constructions in natural language which make use of the notion of a motivating reason — we can call them **reasons contexts** — are factive. Here, for example, are Maria Alvarez, Peter Unger, and Timothy Williamson (respectively).¹¹

The expression 'N's reason is/was that' is an operator that, it seems, can only form true sentences when attached to a true sentence... For, as we saw, there is an implicit

¹⁰ See Alvarez [2010: 42] and Hyman [2011: 363-364], respectively, for the following to quotes. See also Williamson [201X: 3, 14-15, 22].

¹¹ See Alvarez [2010: 136-137], Unger [1975: 208], and Williamson [201X: 30n21], respectively. The redacted piece of Williamson quote continues, "projective talk of the sort mentioned in fn. 10;" footnote 10 discusses "projective" talk of knowledge which suggests that knowledge is not factive.

contradiction in claims to the effect that someone's reason is (or was) that p but not p , that is, claims such as 'Alice's reason was that her husband was at home, although he wasn't at home.'

It is inconsistent to say "His reason was that the store was going to close, but it wasn't going to close."

Philosophers often seem to misunderstand the bearing of 'reasons' through a literal-minded treatment of projective talk of the sort [which occasionally misleads people into denying that knowledge entails truth].

The thread of thought here seems to be that nonfactive uses of phrases like "A's reason is that r " and "Her reason for Φ ing was that r " are somehow inappropriate. Indeed, they are described with notions such as *contradiction* and *inconsistency*; comparisons to claims of the form " A knows that p " are made; Moore's paradox is even invoked.¹² Furthermore, these characterizations are given without argument, leaving one to feel that they are supposed to be self-evident.

It is far from being the case that all ears hear these reasons contexts as factive. But the discussion need not devolve into a stubborn debate over whose ears are more discerning of the truth. We have seen that the initial conceptual analysis and casuistry is at odds with this way of thinking. For if there were really a strict inconsistency or incoherence in claims of the form " A 's reason is that r , but it is not the case that r ," we would expect the conceptual analysis of either a *motivating reason* or *acting for a reason* to bear that out. We would expect to find the concept of a fact somewhere internal to the concept of a reason. Here is John Hyman echoing this sentiment:

Let us say that a sentence-forming operator O on one more more declarative sentences is factive if, and only if, the statement " $Os_1 \dots s_n$ " cannot be true unless the statements " s_1 " and " s_n " are true. If this is what factivity is understood to be, it is not just a brute fact about language. Dancy says, "Knowledge is factive, because 'he knows that it is raining but it is not' is uninterpretable." But this gets things the wrong way round. "He knows that it is raining but it is not" is uninterpretable—

¹² See Alvarez [2010: 134] for the supposed analogy to Moore's paradox.

or cannot be true—because “he knows that” is factive; “he knows that” is factive because one cannot falsely or erroneously know that something is the case; and one cannot falsely or erroneously know that something is the case because knowing that something is the case is a relation between knowers and facts or truths.¹³

I take it one of the points here is that, roughly, Moore-paradoxical sentences are paradoxical in virtue of the proper analysis of the concepts they invoke. The reason that “*A* knows that *p*, but not *p*” is paradoxical, Hyman seems to say, is that it employs a factive *concept* (knowledge) nonfactively. That explanation is not readily available, however, when searching for an account of the alleged paradoxical nature of claims of the form, “His reason is that *r*, but it is not the case that *r*.” When investigating the concept of a reason, one finds concepts such as *consideration taken to favor*, *premise of practical reasoning*, and *consideration which persuaded an agent to act* — none of which include, imply, or even hint at the concept of a fact.

We might on these grounds content ourselves to follow the conceptual analysis, and discount the supposed evidence provided by the linguistic sensitivities of these philosophers. Lest we give this opposing line of thought short shrift, though, I think it is worth considering the idea that *in natural language*, reasons contexts are factive — despite the evidence that the concept of a reason itself is not factive. That is: it is possible that our usage of the term ‘reason’ departs from the proper conceptual analysis of the notion of a reason. This is perhaps friendly to Williamson’s diagnosis concerning nonfactive, “projective” uses of the word ‘knowledge:’ the conceptual analysis and the natural language usage of knowledge simply come apart. It is worth noting, though, that Williamson’s inclination, like Hyman’s, seems to be to trust what is suggested by the conceptual analysis, and discount the deviant, nonfactive, natural language usages of ‘knows’ as misleading. So, even if we did find that reasons contexts are factive, there

¹³ See Hyman [2011: 358-359].

would be a serious question as to why this finding should outweigh, when asking whether motivating reasons are facts, what seems to be true of the concept of a reason.

Still, given that our linguistic practices often (rightly, we think) inform our conceptual analysis, perhaps there is something to be learned from further analyzing reasons contexts. The question is: are reasons contexts factive? The answer, I think, is negative. To see why, consider first the intuitive similarity between these two sentences:¹⁴

(5) John's belief was that P.

(33b) Oscar's secret desire is that he be well liked.

Semantically, these two sentences appear to be isomorphic. Here is James Pryor on how these two claims should be semantically analyzed:

In both (5) and (33b), we should understand the that-clause to be supplying an argument to the cognitive nominal that precedes the copula.¹⁵

This sentence appears, semantically, as though it should receive the same treatment as (5) and (33b). But if that is so, there is an important consequence for what we can say about the semantics of that-clauses in reasons contexts. Here is Pryor once more:

The intuitive similarity between (5) and (7) suggests that something analogous is going on with (7)... That is, even if the orthodox semantic analysis of that-clauses is right, and "that P" is a singular term, the role it's playing in (7) is to *specify* John's reason, not to designate something that (7) *equates* with John's reason. Thus, (7) isn't offering us any direct insight into the ontology of reasons.

Now, the intuitive semantic similarity between (5) and (7), combined with the obvious fact that (5) is not a factive construction, might seem sufficient evidence that

¹⁴ In what follows, I draw heavily on the work of James Pryor [2007].

¹⁵ See Pryor [2007: 240] for both this quote and the next. Alvarez [2010: 40ff] seems to recognize that we can move back and forth between purely nominal expressions of reasons and propositional expressions of reasons.

claims like (7) are also not factive.¹⁶ While that seems to me to be correct, I think that Pryor’s insight here runs deeper. What Pryor points out is that that-clauses in reasons contexts seem only to *specify* the reasons in question, the same way that that-clauses in belief contexts specify the contents of beliefs. If John believes that Mary is ill, and it is true that John believes that Mary is ill, neither the truth nor the assertion of this sentence do anything to establish that John’s belief — the mental state *itself* — is identical to the proposition that Mary is ill (though the *content* of John’s belief may be identical to something like the proposition that Mary is ill). In fact, neither the truth nor the assertion of this sentence seem to tell us anything at all about the ontology of John’s belief. The semantic similarity of belief contexts and reasons contexts should lead us to say the same thing about that-clauses in reasons contexts.

We so far have no reason — aside from the linguistic sensitivities of some Factivists — to suspect that reasons contexts should be factive. But if Pryor is right (and I know of no objection to his analysis), then even if that-clauses in reasons contexts could only be properly completed by truths, this would still not reveal anything about the *ontology* of reasons. Even if phrases of the form, “*A*’s reason was that *r*” could only be true in natural language if *r* is true, it would be a mistake to take this as evidence that motivating reasons fall into any particular ontological category (such as facts — or even propositions, for that matter). The ontology of reasons, of course, is precisely what the disagreement between the Factive View_M and the Nonfactive View_M concerns. The linguistic data simply provide us with little evidence to settle this disagreement.

If all of this is correct, those who proceed from the observation that phrases of the form “*A*’s reason is that *r*” are factive may be without much to stand on other than their

¹⁶ Alvarez [2010: 137-138] seems unimpressed by the comparison between belief contexts and reasons contexts. But her treatment includes little mention of the semantic similarities between the two contexts, and she gives little independent reason for thinking that the comparison is unwarranted or misleading.

own linguistic sensibilities. In the spirit of charity, though, perhaps we might try to offer an account as to why their seems to some to be an “air of paradox” — as Alvarez puts it — about claims like (for example) “My reason for giving him the money is that he needs it, although he doesn’t.” I believe that there is a satisfying explanation of the awkwardness of such claims, which does not appeal to the idea that reasons contexts are factive. The explanation is that, as I will argue in the next chapter, claims of the form “ A Φ s for the reason that r ” and “ A ’s reason is that r ” imply that A believes that r . It makes no sense to say, “My reason is that he needs the money, but I do not believe that he needs the money.” This is because, as I will argue, acting for a reason requires believing that reason to be the case.¹⁷ Furthermore, in the spirit of Moore, we can note that it is at least puzzling to say things such as, “I believe that he needs the money, but he does not need the money.” Combining these two insights provides an easy explanation of the alleged “air of paradox” about nonfactive uses of reasons constructions. When someone says, “My reason is that he needs the money,” “I believe that he needs the money” is implied; but when someone says, “He does not need the money,” “I do not believe that he needs the money” is implied. So, the assertion that “My reason is that he needs the money, although he does not need the money” seems to imply an inconsistent report as to what the speaker believes.

This account can also explain, I believe, Alvarez’s astute observation that

When an agent insists that a belief of hers that others claim to be false was, nonetheless, her reason for acting, she will typically do so because she does not accept that the belief is false... if an agent still insists that what she accepts to be a false belief is her *reason* we’ll have to wonder what she means by ‘a reason’. For... [there is a] need to retract one’s claim that one’s reason for Φ -ing was that p on being confronted with the fact that not p .¹⁸

¹⁷ I take this to be a point of widespread agreement among theorists of reasons. See Sandis [2013: 41] for someone who notices, in particular, that claims like these are paradoxical.

¹⁸ See Alvarez [2010: 138].

If I tell you that “My reason for doing it is that he needs the money,” and you confront me with (read: *convince me of*) the fact that he does not need the money, it would indeed be puzzling if I continued to assert that I am giving him money for the reason that he needs the money. But that puzzlement can be explained in exactly the way just illustrated. It would be odd for me to say, “Okay — you have convinced me that he does not need the money; still, I am going to do it for the reason that he needs the money.” Such a sentence, due to the connection between acting for a reason and believing that reason, seems to imply both that I believe that he needs the money and that I do not believe that he needs the money. *That* — the implied inconsistent belief reports — is the source of the oddity, and also of the pressure to retract my report about what my reason is.

Note that there is not much of a puzzle, on the other hand, if I reply, “Fair enough — I see why you think he does not need the money. But I still believe that he needs the money; and that he needs the money is my reason for giving him the money.” The Factive View_M seems to make the wrong prediction about this variation of the case: it seems to predict that my claim that my reason is that he needs the money remains paradoxical in this case. For, according to the line of thought under consideration, *any* sentence of the form “*A*’s reason is that *r*, but it is not the case that *r*” is paradoxical. What we know about *A*’s beliefs should make no difference to the paradoxical nature of such claims. But it is clear that, in the variation of the case at hand, the knowledge of what I believe *does* dissolve (or at least mitigate) the paradoxical nature of my reason-report.

We can make a similar observation about the difference between reasons contexts created from the first-person as compared with the third-person, and also those constructed in the present tense as compared with those constructed in past tense.¹⁹ For

¹⁹ See Alvarez [2010: 134n13] for the claim that the first-person and third-person cases are similarly paradoxical.

my part, there is not much paradox about saying, “His reason is that he needs the money — although, as of course we know, he does not in fact need the money.” But even if this third-person construction were paradoxical, it is surely *less* paradoxical than the assertion that “My reason is that he needs the money, although he does not.” The degree of puzzlement also seems to vary with tense. If we switch from the present tense to the past tense, the air paradox seems to disperse, if not vanish completely. There is no real puzzle about saying, “My reason for giving him the money was that he needed the money. Of course, I now see that he did not need the money; but at the time, I thought that he did.” One could stubbornly insist that there is something uninterpretable about such a claim (although I can see no good reason to insist so); but surely everyone must concede that the nonfactive use of ‘reason’ is *less* puzzling in the past tense than in the present tense. However, these differences created by perspective and tense are not phenomena that the Factive View_M can easily countenance; the Factive View_M seems to predict that all claims of the form “A’s reason is/was that *r*, but not *r*,” are equally paradoxical. On the other hand, my proposed account on which we explain the apparent puzzles by appealing to the beliefs of agents has no such trouble. The first-person cases are more puzzling because the assertion that, “My reason is that *r*, but not *r*” seems to imply the claim that “I believe that *r* and I do not believe that *r*.” The claim that “His reason is that *r*, but not *r*” creates no such implication; it only seems to imply, “He believes that *r*, but I do not believe that *r*.” A similar account can be given to explain why the past-tense constructions are at least less puzzling than the present-tense constructions. When I say, “My reason *was* that *r*, but not *r*,” I imply only that “I *believed* that *r*, but I do not *believe* that *r*.”

So it seems we do not, after all, need to appeal to the idea that reasons contexts are factive to explain any of these apparent puzzles. We need only appeal to something like Moore’s original paradox, and the idea that acting for a reason requires believing that

reason. Indeed, this explanation does not only handle well the apparent paradoxes pointed out by some Factivists. It also provides a satisfying account as to why some nonfactive uses of ‘reason’ are not paradoxical (or, at least, less paradoxical than others); whereas the Factive View_M predicts that all nonfactive uses of ‘reason’ should be equally puzzling. We may admit that there is an air of paradox about some nonfactive uses of reasons contexts — but we should ultimately say that it is only an air, blown in by the genuinely paradoxical nature of saying both that “*A* believes that *p*” and that “*A* does not believe that *p*.”

Summary Remarks

If all of this is correct, the default view concerning the factivity of both acting for a reason and motivating reasons should be the Nonfactive View_M. We find in the conceptual analysis of these notions no hint of the concept of a fact; and we even find some evidence that these notions *cannot* be factive. Still, there is the suggestion that natural language constructions of the form “*A* Φs for the reason that *r*” and “*A*’s reason is that *r*” cannot be true unless *r* is true. But this suggestion turns out not to have much force, since: (i) the conceptual analysis does not bear out the idea that these constructions are factive; (ii) the semantic analysis does not bear out the idea that these constructions are factive; and (iii) even if the semantic analysis did suggest that these constructions are factive, this would not entitle us to any inferences about the ontology of reasons. It is sometimes suggested that the factivity of these constructions would explain the air of paradox surrounding their nonfactive uses, and also the felt need to retract reasons claims when it is demonstrated that the thing claimed to be a reason is something that is not the case. But these things can be just as easily, if not more satisfactorily, explained by the idea — to which all parties will agree — that acting for a reason implies believing that reason.

Furthermore, this connection between reasons and beliefs seems to give satisfying accounts of other cases in which the Factive View_M simply gives the wrong prediction. As I will show in the next chapter, this connection between reasons and beliefs can be easily accommodated by the Nonfactive View_M.

If the Factive View_M should be ultimately favored, then, it will have to find its support from some other line of argument. I devote the rest of this chapter to discussing a formidable candidate for such a line of argument.

REASONS, EXPLANATIONS, AND FACTIVITY

Consider now a variation of the example discussed at the beginning of the chapter.²⁰ Imagine a pond that has thin ice in the middle. Edna takes it that the ice in the middle of the pond is thin. So, when she skates, Edna keeps to the edge of the pond. Now, imagine a pond with thick ice throughout. Edmund takes it that the ice in the middle of the pond is thin. So, when he skates, Edmund keeps to the edge of the pond.

We might ask *why* Edna and Edmund are skating near the edge of the pond. Insofar as we take ourselves, in asking this question, to be asking *for what reason* Edna and Edmund act as they do in skating near the edge, we express some sympathy for the claim that

²⁰ I am adapting this example from Hornsby [2008: 251], which to my knowledge traces back to Gilbert Ryle [1949: 117-118] — although its more contemporary usage may be due to Unger [1975: 209-211]. I have changed some of the details of the case in order to avoid the epistemological baggage that comes with Hornsby's version.

REASONS-EXPLANATIONS: a motivating reason is the kind of thing that can play a role in explaining an agent's action (when the action is one the agent does for a reason).²¹

The idea that motivating reasons play an explanatory role is extremely plausible. Kieran Setiya has even gone so far as to suggest that it is conceptual truth: “the sense of “reason” in which one takes something as one’s reason is *explanatory*”.²² The very concept of a motivating reason might seem to be one that picks out entities with explanatory power. One way to see this is to revisit the supposition that when we ask *why* Edna and Edmund are doing what they are doing, we are asking *for what reason* they are doing what they are doing. When we ask “Why?” questions about actions, we usually take answers to these questions to be giving explanations of the actions in question. This indicates a quite intimate connection between motivating reasons and action explanations — even if the intimacy does not approach the level of identity, or amount to a conceptual truth. Indeed, the apparent intimacy of the connection here might lead one to hope for action explanations to be given in terms of, or wholly exhausted by, reasons for which agents act. One might also hope, in the present case of Edna and Edmund, for two different explanations of what the two different agents are doing. This hope is founded on the observation that there is a seemingly significant difference between Edna and Edmund: Edna has a true belief about the ice, while Edmund has a false belief. To the extent that we expect this feature to make a difference to the explanations of their actions, we express some sympathy for the claim that

²¹ Sandis [2013: 30] rightly refers to this claim as a common “assumption” in the literature. Setiya [2007: 39-47] provides a nice defense of the axiom if one is needed. The earliest contemporary treatment of something approximating this thought, to my knowledge, is found in G. E. M. Anscombe [1957: 9ff].

²² See Setiya [2007: 42], and Setiya [2007: 23] for the same sentiment redescribed.

FACTIVE EXPLANATIONS: something that is not the case cannot explain anything.²³

This claim is also very plausible — even if, unlike REASONS-EXPLANATIONS, it does not seem to be a conceptual truth. Consider a paradigm class of explanations: causal explanations. Suppose your office is adjacent to mine, and while you are working, you hear what sounds like glass breaking in my office. You rush into my office and notice that one of my windows is broken. You proceed to ask me why the window is broken, and I respond with something like: “My window is broken because a baseball was thrown at it.” It would be quite puzzling if, when you go on to ask me where the baseball is, I respond, “Oh, there is no baseball — it is not the case that a baseball was thrown at my window.” You should like to know, I think, just what I mean by “because.” My “because” claim seems to make no sense; claims of the form “*P* because *Q*” (and perhaps, “What explains *P* is that *Q*”) seem to be doubly factive: they can be true only if both *P* and *Q* are true.²⁴ Furthermore, since claims of this form seem to be paradigmatic explanatory claims, we might take their factivity to be good indication that explanatory claims in general are factive.

So far, I have been describing a natural way of looking at actions and explanations of them. But anyone looking at things in this way might feel disappointed by the natural way in which we were looking at acting for a reason earlier — the natural way of looking at things which led us to believe that the Nonfactive View_M is the intuitive view of acting

²³ As with the previous axiom, this is a widely held assumption. But it is not difficult to find statements of the assumption, as in Dancy [2014: 83], Hyman [1999: 443], Littlejohn [2012: 102] Sandis [2013: 31, 46], Scott Sehon [2005: 178], P. F. Strawson [1992: 109], and Schnee [2015: 1].

²⁴ I have borrowed the phrase “doubly factive” from Dancy [2014: 84].

for a reason. The disappointment is likely to come when inquiring about pairs of agents like Edna and Edmund. The case of Edna is straightforward enough for the Factivist and Nonfactivist alike. The natural answer to the question, “Why is Edna skating near the edge of the pond?” is that thin ice: the ice in the middle of the pond is thin. This explanation respects both of the intuitive claims we just observed: it explains the action via a reason for acting, and it provides as the explanation something that is the case. But the same explanation is not available in the case of Edmund. We certainly cannot answer the question, “Why is Edmund skating near the edge of the pond?” by saying that the ice in the middle of the pond is thin.²⁵ Since it is not the case that the ice in the middle of the pond is thin, such an explanation (aside from being plainly incorrect *to say*) strictly violates our commitment to FACTIVE EXPLANATIONS. Notice, though, that the Nonfactive View_M allows that Edmund’s reason for acting could be that THIN ICE even though it is not the case that thin ice. Furthermore, we have agreed by our commitment to REASONS-EXPLANATIONS that Edmund’s reason for acting — whatever it may be — is the kind of thing that can play an explanatory role vis-à-vis his action. So it seems that the Nonfactive View_M is set up to allow for a conflict between REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS; and those who endorse the view will be forced, on pain of contradiction, to deny at least one of these very plausible claims.

So there seems to be a real puzzle here about actions, reasons, and explanations. Intuitively, all three of the Nonfactive View_M, REASONS-EXPLANATIONS, and FACTIVE EXPLANATIONS are incredibly plausible; but they also seem to be jointly inconsistent. The puzzle has the potential to be a powerful criticism against the Nonfactive View_M. The

²⁵ For the remainder of this manuscript, I will sometimes use THIN ICE as a convenient shorthand (and I will use other similar shorthands with different contents). As it makes no difference to my discussion, I make no distinction between the *proposition* that the ice in the middle of the pond is thin and the *state of affairs* of the pond having thin ice in the middle. Neither do I make any distinction between the *true* proposition, the *obtaining* state of affairs, or the *fact*. In similar spirit, I often use “false” and “not the case” interchangeably.

puzzle revolves around a certain kind of case: cases like that of Edmund, wherein the agent is wrong about how the world is, and then acts in the light of that erroneous view of the world. Call such cases **error cases**.²⁶ The criticism, very simply put, is that the Nonfactive View_M is not compatible with correct and satisfying reasons-explanations of actions in error cases, and it seems to provide no theoretical tools for avoiding obviously problematic explanations in error cases. If these difficulties were unavoidable in general, we theorists might begin considering whether we should do away with one of either REASONS-EXPLANATIONS or FACTIVE EXPLANATIONS. However, the combination of these two axioms with the Factive View_M appears to create no such problems in error cases. According to the Factive View_M , A can Φ for the reason that r only if it is the case that r . Since it is not the case that the ice in the middle of the pond is thin, then, Edmund could not have acted for that reason. Since Edmund could not have acting for that reason, there is no temptation to answer the relevant “Why?” question by citing that false consideration as his reason. There is therefore no looming conflict, given the truth of the Factive View_M , between REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS. Indeed, perhaps what all of this shows is that the Nonfactive View_M is ultimately a nonstarter despite its initial plausibility; and we theorists should direct our efforts toward developing the best version of the Factive View_M possible.

THE ARGUMENT FROM THE FACTIVITY OF EXPLANATION

Or so one chief argument against the Nonfactive View_M goes. Now that we have a good idea of what could motivate critics of the Nonfactive View_M to raise this objection, we can take a closer look at it. But first, I engage in an ultimately orthogonal, but nonetheless necessary, aside concerning the dialectic surrounding this argument.

²⁶ This is now a standard way of labeling such cases, but see Alvarez [2010: 50, 124; 2016].

I have sometimes been met with puzzlement when suggesting that error cases should present a problem for the Nonfactive View_M. Indeed, some have suggested to me that such cases appear to present a problem for the Factive View_M. Such puzzlement, I think, has two primary sources. One source is that, insofar as it seems true that agents *can* act in the light of a falsehood (that is: they can Φ for the reason that r even if it is not the case that r), the conceivability of cases in which agents act in the light of falsehoods poses a putative problem for the Factive View_M, since it holds that there are no such cases. As I see it, this amounts to not much more than placing an intuitive burden on the Factive View_M for giving the wrong verdicts in some cases. The problem under consideration, though, is not just the task of saying whether it is possible for agents to act in the light of falsehoods. The task is reconciling that possibility with the ideas that (i) reasons have explanatory power and (ii) explanation is factive.

A second source of confusion here is the thought that error cases create an argument not for the Factive View_M, but instead for the view often called Psychologism. This is the view that, roughly, all motivating reasons are psychological facts about the acting agents. Non-Psychologism (or “Anti-Psychologism”) is the denial of Psychologism. Here is James Lenman, describing how error cases are supposed to benefit Psychologism:

The biggest headache for anti-psychologists such as Dancy however is furnished by cases where the agent’s belief is *false*. The fact of Angus’ being fired is naturally adduced to explain his punching his boss in cases where he has indeed been fired. But in cases where Angus punches his boss, believing mistakenly that he has been fired, it seems quite wrong to say he so acts because he has been fired. In such a case we surely must retreat to a psychologised explanation if we are to have a credible motivating reason explanation at all.²⁷

²⁷ See Lenman [2011: 20-21]. See Alvarez [2016: 4-5, 24ff] for a similar characterization of the dialectic; but see Sandis [2013: 32] for an example of one who seems to recognize that the real issue here is that of factivity. A fuller discussion of Psychologism comes in the next two chapters.

The crux of the argument comes at the end. Psychologism is thought to benefit from such cases because, in such cases, it is tempting to think that the only available explanation is the psychologized one: Angus punched his boss *because he believed* that he was fired. Whence this temptation, though? As Lenman’s analysis itself reveals, the temptation is brought about by the thought that we cannot explain Angus’ action by simply giving the content of (what seems to be) his reason for acting (“he was fired”). But *that* thought is appealing only because the consideration that initially seems to be Angus’ motivating reason is a falsehood. To put it in the terms we have used so far: explaining Angus’ action by appeal to this false reason would seem to violate the axiom of FACTIVE EXPLANATIONS. So the pressure to turn to Psychologism in such cases ultimately has nothing to do with Psychologism and Non-Psychologism *per se*. The driving force of the case is found in the tension between saying that (i) Angus acted for the reason that *r*, although not *r*; (ii) Angus’ motivating reason can serve to explain his punching his boss; and (iii) only facts can serve to explain Angus’ punching his boss. This is just an instance of our more general present problem: that of reconciling the Nonfactive View_M with REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS. Psychologism seems appealing in such cases only because it dissolves this trilemma, by saying that motivating reasons are psychological *facts* (and therefore, giving up the Nonfactive View_M). So, the argument for Psychologism here — if there is one at all — ultimately proceeds by way of a more fundamental argument for the Factive View_M.²⁸

With these confusions out of the way, we may proceed to the argument itself, which I will refer to as the **Argument from the Factivity of Explanation** (or,

²⁸ One might conjecture that the confusion is compounded by the fact that many Non-Psychologists are at pains to argue, against Psychologism, that motivating reasons are *facts* rather than psychological *states* of agents. Since error cases are often thought to present a problem for Non-Psychologists, this fact may also naturally (though erroneously) lead to the thought that error cases present a problem for Factivists.

“AFE”). The argument appears in many forms in the literature. These vary in the extent to which they explicitly state that the nonfactivity of motivating reasons is at odds with motivating reasons playing an explanatory role; but the sentiment has nonetheless more than worked its way into the air.²⁹ Here is Dancy’s rendering of AFE:

This Edmund example faces us with an inconsistent triad.

1. We explain an action by giving the reason for which it was done.
2. Edmund’s reason for keeping to the edge of the pond was that (as he had been told) the ice in the middle was thin.
3. That the ice in the middle was thin is not (any part of) what explains Edmund’s keeping to the edge.

Here also is Pamela Hieronymi:

If Erin’s operative reason for leaving was that the meeting was over, but the meeting was not over, then we cannot appeal to the meeting’s end to explain her departure—because the meeting did not end. Something that is not the case cannot explain something that is. To provide an explanation of Erin’s departure, one must cite some fact.

Finally, for one more example, here is Michael Smith:

But if [John] *failed* to illuminate the room then he would be forced to say that the belief is what moved him, as in that case it wouldn’t be true that flicking the switch illuminated the room. Quite in general, then, when we are motivated by a false view of a situation, it is not true that what moves us is... something about the nature of the situation... Only explanatory connections that go via the psychologies of the agents whose bodies move are able to explain their actions. So even though Dancy says that what motivates us is the nature of the situation, it seems that even he has to admit that which features of a situation motivate us is a function of our psychology.

Abstracting away from some of the particularities of each of these, it is perhaps natural to put the argument in the form of a trilemma:

[_] All motivating reasons are explanatory;

[_] All things that are explanatory are facts;

²⁹ See Dancy [2014: 83], Hieronymi [2011: 410-411], and Smith [2012: 391-392], respectively.

[_] Some motivating reasons are not facts.

Since we are looking for a positive argument in favor of the Factive View_M, though, it might be more straightforward to put the argument this way:

[1] All motivating reasons are explanatory;

[2] All things that are explanatory are facts;

[3] Therefore, all motivating reasons are facts.

Here claims [1] and [2] are articulations of REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS, respectively. Of course, if the argument is sound, and claim [3] is true, then the Factive View_M is true and the Nonfactive View_M is false.

In this section I want to begin discussing what I take to be a major shortcoming of this argument. Indeed, I think the shortcoming turns out to be fatal. The shortcoming is that both claims [1] and [2] are ambiguous; their ambiguity is owed to an ambiguity in the notion of *being explanatory*. What does it mean for something — a motivating reason, for example — to be explanatory? Of particular interest to present purposes is the question: what does it mean for a reason to be explanatory with respect to an action done for a reason? What does a claim like [1] amount to? Here are some options from the literature.³⁰

³⁰ See Anscombe [1957: 9], Hornsby [2008: 250], Alvarez [2010: 170], Dancy [2014: 83], and Sandis [2013: 32] (respectively), for the following passages. I expect claim [1a] below will be agreeable to most everyone in the area; see also Peter Achinstein [1983], Dancy [2014: 90], Wayne Davis [2003: 456f], and Hieronymi [2011: 409]. For claim [1b], see also Hyman [1999: 443; 2011: 359], McDowell [2013: 18-19], and Sandis [2013: 32]. For [1c] I also draw on Carl Hempel and Paul Oppenheim [1948: 137ff]. I expect there would be wide endorsement around [1c], save for Sandis [2013: 33] and perhaps Dancy [2014]. For [1d] see also Hyman [1999: 443]; but again I expect all in the area would agree on something at least as strong as this. I expect again that [1e] would be agreeable to most in the area; it was suggested to me by Maria Alvarez.

[Intention actions] are the actions to which a certain sense of the question ‘Why?’ is given application; the sense is of course that in which the answer, if positive, gives a reason for acting... the question “What is the relevant sense of the question ‘Why’” and “What is meant by ‘reason for acting?’” are one and the same.

X Φ -d because p (where ‘because’ can be glossed with ‘for the reason that’)... where there is a reasons-explanation from X’s knowledge, the fact that *p* was a reason X had for Φ -ing.

A reason explanation, in the sense in which I am using this phrase, is an explanation whose explanans is the agent’s reason for acting.

Every statement we make, in giving an explanation of anything, needs to be true. But explanations often contain clauses that are themselves capable of truth and falsity, and if the truth of the whole does not depend on the truth of the contained party, we think of the context as intensional.

Even in cases where we... explain her action with a simple ‘she did it because *p*’ what does the explaining is not the falsehood ‘that *p*’... The reason cited does not itself even contribute to the explanation.

From these we can perhaps glean an initial menu of options for disambiguating claim [1]:

[1a] A reason is the kind of thing that can be given in an answer to a “Why?” question.

[1b] A reason is the kind of thing that can replace *q* in a claim of the form “*P* because *q*”.

[1c] A reason is the kind of thing that can be the *explanans* of an action explanation.

[1d] A reason is the kind of thing that can be a part of an action explanation.

[1e] A reason is the kind of thing that can contribute to the explanation of an action.

I want to start by making a few observations about this menu of options. First, it seems to me that none of these disambiguations of claim [1] can lay claim, pre-theoretically, to being *what it really means* for a reason to be explanatory with respect to an action. At least initially, each of these claims seems to adequately, though differently, express the idea. But if that is right, AFE runs into a complicating question right from the start: which of these disambiguations is being used in the argument for the Factive View_M? On the heels of this observation is a second one: advocates of AFE must choose (and perhaps defend) some particular disambiguation to run the argument on, lest the opponent dodge the argument by simply claiming that she meant something else by “explanatory.” A third observation here is that, if any of these disambiguations is substituted into our current version of AFE, the argument will be rendered invalid. This may be obvious, but here is a quick demonstration using the first disambiguation:

[1a] A reason is the kind of thing that can be given in an answer to a “Why?” question.

[2] All things that are explanatory are facts;

[3] Therefore, all motivating reasons are facts.

Claim [3] clearly does not follow from claims [1a] and [2]. In order to produce a valid argument using [1a], claim [2] will have to be revised as well. This will be a complication when any disambiguation of claim [1] is substituted into AFE.

Finally, I want to note that even these five disambiguations seem to admit of further disambiguation — and they do so in a manner that is suggestive of more general

ways of understanding the notion of *being explanatory*.³¹ For example, claims [1a] and [1b] seem to associate the notion of explanation with a human practice, often carried out in conversation. We ask questions of a certain kind (or form), and we try to give answers of a certain kind (or form). The explanatory nature of reasons might be understood according to the role reasons (or their contents) play in this practice. Claims [1c], [1d], and [1e] seem to abstract away from some of these particularities. They seem to associate the notion of being explanatory with a certain kind of proposition, sentence, or assertion — or perhaps with a certain kind of relation between propositions, sentences, or assertions. More general distinctions such as these may be useful to keep in mind as our examination of AFE proceeds.

‘Why?’ and ‘Because’

With these preliminaries in mind, we can return to the complication that occurs when one tries to substitute claim [1a] in for claim [1] in AFE. The resulting argument is invalid. Fair enough, one might think — but it is obvious enough how to revise the second premise and fix the argument here:

[1a] A reason is the kind of thing that can be given in an answer to a “Why?” question.

[2a] Anything that can be given in an answer to a “Why?” question is a fact.

[3a] Therefore, all reasons are facts.

³¹ Thanks to Dan Bonevac for helping me to clarify my thinking with these categorizations. The reader will see that things could have been grouped in more than one way, since many of my critiques will cut across categories.

Recalling how we started our inquiry wondering about Edna and Edmund, one might suppose that any answer to a “Why?” question counts as an explanation. According to the thought expressed by [1a], one might therefore suppose that anything given in an answer to a “Why?” question also counts as explanatory. So, anything given in an answer to the question, “Why is she Φ ing?” will count as explanatory vis-à-vis her Φ ing.

Much of that seems sensible enough. But to use this conception of explanation in a successful variation of AFE, one must also make use of claim [2a]; and claim [2a] is clearly false. In fact, as far as I can tell, although many theorists would endorse [1a], no theorist would endorse [2a]. This makes this sample argument a useful starting point for us: I expect everyone will agree that the argument is a poor one, and that its obvious shortcoming is in [2a]; and yet this argument provides a model for illustrating the general kind of trouble that AFE faces. To see this, consider the following answer to the question of why Edmund is skating near the pond:

He believes that the ice in the middle of the pond is thin.

This kind of explanation (which I will refer to as **psychologized**) seems perfectly correct, and it seems to me that everyone acknowledges the plausibility of such psychologized explanations. Indeed, as we saw earlier, such explanations are favored by Factivists in error cases like that of Edmund. For, although it is not the case that the ice in the middle of the pond is thin, it is the case that Edmund believes it to be thin. So, in error cases, Factivists will often switch to a psychologized explanation in order to preserve the factivity of the explanation.³² The result of this switch, though, is that explanations in error cases

³² As Dancy [2014: 83] points out, this move is almost always accompanied by a revised account of what the agent’s reason *is*: many Factivists adopt the view that Edmund’s reason is *that he believes* that the ice in the middle is thin. One notable exception to this is Alvarez [2010].

will always proceed by way of falsehood. In the case of Edmund, the falsehood is that the ice in the middle of the pond is thin. If claim [2a] were true, we should expect falsehoods to be prohibited from entering into explanations. So, psychologized explanations render this variation of AFE ineffective twice over. First, claim [2a] is shown, by the plausibility of psychologized explanations in error cases, to be simply implausible. Second, [2a] is shown, by the state of the dialectic surrounding error cases, to be unacceptable by many Factivists' own lights.

In view of this difficulty, advocates of AFE might want to make an adjustment. They might try using, instead of claim [1a],

[1a*] A reason is the kind of thing that can be given *as* an answer to a “Why?” question.

This revision conceives of the explanatory role of reasons not as consisting in just being given *in* an answer to a “Why?” question, but in *being* the answer to a “Why?” question. Now, without any restrictions on what may count as an answer to a “Why?” question, [1a*] will not serve to improve AFE. For one answer to the question “Why is Edmund skating near the edge of the pond?” might be, “His reason is that the ice in the middle of the pond is thin.” If this is an acceptable answer, [1a*] hurts, rather than helps, this argument for the Factive View_M. Any helpful parameter as to what counts as an answer to a “Why?” question, of course, will be one that pressures us to think that the agent's reason must be a fact. Perhaps conception [1b] can be of help here. After all, when we ask questions of the form, “Why did *A* Φ ?” we often expect an answer of the form “Because *q*,” and take *q* to *be* the answer to our question. It might be thought, then, that the explanatory role of motivating reasons consists in being able to replace *Q* in a claim of

the form “ P because q ” (or “What explains p is that q ,” or perhaps, “ Q is the explanation of p ”).

Of course, in order to make use of these ideas in AFE, we must also revise the second premise of the argument. Here is what the argument using [1b] would look like.

[1b] A reason is the kind of thing that can replace q in a claim of the form “ P because q ”.

[2b] Anything that can replace q in a claim of the form “ P because q ” is a fact.

[3b] Therefore, all reasons are facts.

As we saw in the previous section, it does seem clearly incorrect to make a claim of the form “ P because q ” where q is a known falsehood. So perhaps we should accept claim [2b]: such claims are indeed doubly factive. The trouble with this argument is rather in what claim [1b] implies that it means for something to be explanatory. It is indeed natural to offer explanations by making claims of the form “ P because q .” Some might even follow Hyman in his proclamation that “reasons can be stated or given; and the canonical form of a sentence stating or giving a person’s reason for doing or believing something is ‘ A Φ ed because p ’.” Constatine Sandis cites Hyman approvingly, even going to far as to say that “all reasons-statements must be translatable (without change of meaning) to any of the standard forms of explanation, on pain of failing to qualify as explanatory.”

I think that both Hyman’s claim about canonical explanatory form, and Sandis’ claim about translatability, are false; and I will turn to them shortly.³³ It is worth first

³³ See Hyman [1999: 443] and Sandis [2013: 38], respectively. I give fuller treatment to these thoughts in chapter three.

mentioning that even if “ P because q ” is the natural (or even canonical) form of reasons-explanations, it is hardly obvious that all reasons-explanations *must* therefore be given (or be capable of being given) in this form. Now, depending on what is involved in something being *canonical*, and whether Sandis’ claim about translatability is true, it may after all be true that any reasons-explanation must be capable of taking the form in question. But, as we will see, there are other explanatory forms that have just as strong a claim to being the canonical form of a reasons-explanation, and Sandis’ translatability claim is implausible.³⁴

Is there any good independent reason to think that reasons-explanations must be given in the form, “ P because q ,” where p is the fact that $A \Phi d$, and q is the content of A ’s reason for Φ ing? I cannot find any. Indeed, there are good reasons not to think this. The first comes from a general skepticism about appeal to natural language here. It is true that, in English, claims of the form “ P because q ” are used to offer explanations, and that we can give reasons-explanations by giving the content of the agent’s reason for q in such claims. But it is a strange suggestion that whether something counts as explanatory, metaphysically, depends on the contingencies of how any natural language is constructed. This is not to say that data from natural language cannot provide us any insight into the nature of explanation. But we should be skeptical, unless we are presented with good argument, of the claim that playing an explanatory role can be reduced to playing a syntactic or semantic role in natural language. Of course, such a reduction is just what seems to be implied by [1b].

A more substantive reason to reject this conception of the explanatory is that, if *any* one explanatory form can lay claim to being the “canonical” reason-giving form, it

³⁴ On the issue of translatability I make only brief remarks here. See chapter three for more.

seems to me that honor must be reserved for the form: “*A*’s reason for Φ ing is *r*.” It is unquestionable that

[1b*] A motivating reason is the kind of thing the content of which can replace *r* in a claim of the form, “*A*’s reason for Φ ing is *r*.”³⁵

Indeed, claim [1b*] might be an analytic truth. So, if we are going to reduce the explanatory role of reasons to any role in natural language, it seems to me that it should first and foremost be one that satisfies [1b*]. Now, perhaps an advocate of AFE might want to assert [1b] in addition to [1b*]. But we now have independent reason not to allow this addition without qualification. As we saw earlier, natural language claims of the form “*A* reason for Φ ing is that *r*” do not appear to be factive; but claims of the form “*P* because *q*” are doubly factive. So, we should not expect that everything that can be substituted in for *r* in a claim of the first form can also be substituted in for *q* in a claim of the second form. Since we have little independent reason to accept [1b], and some good independent reason to accept [1b*], we should retain [1b*] if we are forced to make a choice. Furthermore, even if there is independent reason to insist on [1b], the comparison with [1b*] casts doubt on the idea that [1b] can be used in an argument for the Factive View_M without begging some question. Since the Nonfactivist presumably holds that reasons contexts are not factive, but ‘because’ contexts are factive, the Nonfactivist should reject out of hand any suggestion that reasons claims are interchangeable with any factive construction. A Factivist aiming to use [1b] to run AFE might try to insist either that they

³⁵ Notice that this scheme does not require that the reason be presented in propositional form. A fuller discussion of this possibility takes place in chapter three. There should be no concern about the switch, since we have already seen that claims of the form “*A*’s reason for Φ ing is *that r*” are not factive; and the propositionality could only serve to increase factive pressure.

ought to be interchangeable or that the because contexts are somehow *primary*, but it is hard to see, in the light of the foregoing considerations, what argument could be offered for these claims.

A third reason to reject the [1b] conception of the explanatory comes from the known success of what I will call **purposive explanations**. Consider the following answers to the question of why Edmund is skating near the edge of the pond:

In order to avoid thin ice in the middle of the pond.

He did it to avoid falling through the ice.

His purpose was to stay clear of thin ice.

These ways of explaining Edmund's action seem perfectly correct, and every bit as natural and familiar as answers of the form "*P* because *q*". Indeed, some might even go so far as to say that it is purposive schemes — not 'because' schemes — that present the canonical form of reasons-explanations of actions. Here are Scott Sehon and G. F. Schueler (respectively):³⁶

[Common-sense] explanations [of action] are not causal; instead, I claim that they are *teleological*. A teleological explanation explains by citing the *purpose* or *goal* of the behavior in question... The paradigmatic form of a teleological explanation is *A* Φd in order to Ψ.

Explanations of actions in terms of the agent's reasons, I want to claim, work by citing the purpose or purposes for which the person who performed the action actually acted.

³⁶ Notice that this scheme does not require that the reason be presented in propositional form. A fuller discussion of this possibility takes place in chapter three. There should be no concern about the switch, since we have already seen that claims of the form "*A*'s reason for Φing is *that r*" are not factive; and the propositionality could only serve to increase factive pressure.

However, to the extent that we accept such explanations, we should resist at least [1b].³⁷ Purposive explanations appear to comport well with [1a*], and even [1b*]. We can satisfactorily answer the question, “Why did A Φ ?” with the answer, “In order to Ψ .” Following the model presented to us by ‘because’ schemes, we might take the ‘in order’ operator to signal that what comes next is being offered up as the bit with explanatory power. So, insofar as we accept these kinds answers as reasons-explanations, we may also be inclined (correctly, I think) to accept claims of the form, “ A ’s reason for Φ ing is to Ψ .” Here we capture the agent’s reason for acting in an infinitival clause. But one can substitute neither the infinitival clause nor the content of that clause for q in a claim of the form “ P because q ”. It makes no sense to say, “He did it because to avoid thin ice,” or, “He did it because stay clear of thin ice.” So the plausibility of purposive explanations appears to make [1b] implausible.

It is possible to push back here by saying that reasons and purposes are very different beasts, ontologically speaking. If reasons and purposes are not to be identified, then perhaps this talk about explanations that appeal to purposes is a red herring, and should not be taken as evidence against [1b] as a claim about the explanatory nature of reasons. I postpone fuller discussion of purposive explanations for later. This resistance can be held off, for now, by noting that explanations of the form, “ A Φ d in order to Ψ ” can be reasons-explanations (and clauses of the form “to Ψ ” can be used to specify an agent’s reason) even if purposes are not motivating reasons. Recalling Davidson’s well-known thought, we can see that these purposive explanations are in the business of doing exactly what reasons-explanations do: they reveal the good the agent saw in doing what she did, and specify the considerations that motivated her to act as she did. Now, this

³⁷ Such schemes are receiving increased attention; Scott Sehon [2005] provides a nice discussion. These same criticisms will be applicable to any version of AFE making use of claim [1c] (despite my categorizing them separately). For the sake of brevity and focus, I do not demonstrate this in the main body of the paper.

revelation and specification comes in a clause of infinitival, rather than propositional, form. One might then quibble about whether the clause, “to avoid thin ice” *really* specifies a motivating *consideration*, since things considerable must be propositional. But surely there is no real confusion here about what motivates, for example, Edmund to skate around the edge of the pond. We could hardly respond to someone who offered us this purposive account, “Yes, that seems right. But *why* is he skating around the edge of the pond? I was hoping to know his reason.” His reason has been given by the thought that he is skating so as to avoid thin ice. So the bottom line, for now, should be just this: if a reasons-explanation can be given by a claim of the form, “to Ψ ,” then claim [1b] looks implausible. It may turn out that reasons-explanations cannot be given in this form, but that remains to be seen.³⁸

The points raised so far, I think, give us good reason to believe that any variation of the Argument from the Factivity of Explanation that relies on precisifying the explanatory role of reasons in the manner of [1a] or [1b] is likely to fail. One simply does not find, in investigating the nature of answers to “Why?” questions and “because” claims, the necessary resources for mounting a successful argument for the Factive View_M.

Explanatory Sentences

In view of these difficulties, advocates of AFE may be wary of putting too much weight on questions and claims of certain forms. There is a way of conceiving of the notion of being explanatory that abstracts away from some of these difficulties. One might think generally of *explananda* and *explanantia* as sentences or sets of sentences (or

³⁸ See chapter three for more on whether purposive explanations count as reasons-explanations.

perhaps propositions or assertions). Here are Hempel and Oppenheim laying out their well-known sentential framework for explanation:

By the explanandum, we understand the sentence describing the phenomenon to be explained (not that phenomenon itself); by the explanans, the class of those sentences which are adduced to account for the phenomenon.³⁹

So, for example, in trying to explain the phenomenon of Edmund's skating near the edge of the pond, we have as our explanatory target the sentence describing that phenomenon. Perhaps it is something like, "Edmund is skating near the edge of the pond". We then search for another sentence (or, perhaps, a set of sentences) that would account for the phenomenon described in this sentence.

This way of thinking immediately calls for further disambiguation of the notion of *being explanatory*. Being explanatory could here mean *being the (whole) explanans* in the sense of being the whole set of sentences which completely explains the phenomenon in question. But it could also mean *being a subset* of those sentences (which, by themselves, give only part of the whole explanation). Zooming in a bit more, being explanatory could mean *being a component* of a sentence: being a crucial that-clause — or perhaps a noun-phrase — in one of the sentences in the set. Finally, being explanatory could even mean being a crucial *referent* of one of the sentences of the whole set. There may be other options here as well.

So any version of AFE that conceives of explanation in this way will have to further clarify the notion of being explanatory. Now, if one holds that being explanatory amounts to being the *explanans* in the sense just described, the one might endorse claim [1c]. The version of AFE operating on this disambiguation would go as follows:

³⁹ This way of thinking about explanation is common, but see Hempel and Oppenheim [1948: 136-137]. Thanks to Dan Bonevac for encouraging me to give fuller consideration to this framework.

[1c] A reason is the kind of thing that can be the *explanans* of an action.

[2c] Anything that can be the *explanans* of an action is a fact.

[3c] Therefore, all reasons are facts.

It is perhaps easy to slide into accepting claim [1c] by way of accepting claim [1b]. This is due to the temptation to think of p and q in a claim of the form “ P because q ” as the *explanandum* and *explanans* (respectively). To the extent that disambiguation [1c] just collapses into disambiguation [1b], any argument featuring it will face the problems pointed out in the previous section. One might also come to conception [1b] through the something like [1a*], with the addition that any (proper) answer to a “Why?” question must be a sentence. This might lead us to think that the *explanans* of an action is the sentence that is the answer to the relevant “Why?” question. However, given that two common ways of answering such “Why?” questions are with clauses of the form “to Ψ ” and “for the reason that r ,” and that such clauses do not appear to be sentences, this addition is at least suspect. Perhaps one would like to argue that such clauses are, in fact, sentences. I will come to this suggestion momentarily. In any case, it remains to be seen whether [1a*] and [1c] might rise and fall together; but we should at least not assume that it is safe to move from [1a*] to anything like the sentential conception of [1c].

Setting aside these complications for the moment, we can see that there is good reason to reject claim [1c] if we are adopting the sentential conception of explanation just laid out. On this conception, when we talk of *the explanans*, we are talking about the whole set of sentences comprising the complete explanation. So, in our current example involving Edmund, the *explanans* is taken to be the whole set of sentences that accounts for Edmund’s skating near the edge of the pond. If that is right, though, we should clearly reject claim [1c] as a conception of the explanatory role of motivating reasons. In looking

for a reasons-explanation of Edmund's action, we are looking in particular for those sentences which, at the very least, somehow indicate or make use of the content of Edmund's reason for acting. The whole set of sentences accounting for the phenomenon captured by the sentence, "Edmund is skating near the edge of the pond" will be quite a large complex, including sentences like, "Ice skating is a cold-weather pastime," "Edmund owns ice skates," and "Edmund chose not to skate with Edna today." It borders on absurdity to think that Edmund's motivating reason for skating near the edge is identical to the whole set of sentences including all of these (and more). It is simply not plausible that each and every one of these sentences is part of what motivated Edmund to skate near the edge of the pond.

Now, as was the case with [1a], I do not think that anyone advancing AFE conceives of the explanatory nature of reasons in this way. I think that the more plausible conception in this family of views is that what it comes to for a reason to be explanatory is for it to be a subset of the sentences in the whole set comprising the *explanans*, or perhaps a *part* of one of those sentences. These conceptions sit more comfortably with the fact that discussions about acting for reasons often proceed by attending to examples focused on individual reasons captured by a single that-clause. If we retain the thought that the (whole) *explanans* is the complete set of sentences accounting for the phenomenon in question, we can think of subsets of sentences (even single sentences), and clauses in the individual sentences, as *parts* of the explanation. So perhaps we are moving here toward conception [1d] of the explanatory role of reasons. Making use of that disambiguation, here is what AFE looks like:

[1d] A reason is the kind of thing that can be a part of the explanation of an action.

[2d] Anything that can be a part of the explanation is a fact.

[3d] Therefore, all reasons are facts.

As we just saw, *being a part* here could mean being a set of sentences, being a single sentence, being a clause in a sentence, or something else. In the rest of this section I will argue that none of these conceptions can be used in a successful variation of AFE.

It is perhaps easiest to start by quickly setting aside the “clausal” conception of *being a part*. On this way of thinking, [2d] seems to suffer all of the basic shortcomings we saw earlier in [2a]. If a clausal conception of [2d] is true, for example, then in error cases we cannot make use of psychologized explanations (since the psychological state adduced to explain the action will have its content represented by a false that-clause). We can clearly explain Edmund’s behavior with the sentences, “He believes that the ice in the middle is thin,” “He is doing it for the reason that the ice in the middle is thin,” and “He is doing it to avoid thin ice in the middle;” but neither “the ice in the middle is thin” nor “to avoid thin ice” are facts. I cannot see how this sort of objection could be overcome; but as I said before, I do not think that anyone advancing AFE would actually endorse this reading of AFE. Instead, I think it more likely that Factivists might adopt the view that a reason is the kind of thing that can be a part of an action explanation, where “be a part” means “be a subset of the whole set comprising the complete explanation” (whether that subset be many sentences or just one). In effect, then, [1d] is saying that anything explanatory (including reasons) must be one of these sentences; and [2d] is saying that all of these sentences must be true sentences. For the sake of clarity, we could alter AFE in the following way.

[1d*] A reason is the kind of thing that can be one of the sentences in the *explanans*.

[2d*] Anything that can be a sentence in an *explanans* is a fact.

[3d*] Therefore, all reasons are facts.

There is, in my view, nothing wrong with saying that reasons can be represented in sentential forms, or that things so represented can have explanatory power. But claim [1d*] implies that in order for something to play an explanatory role, it must be capable of being a sentence. At the very least, this requires argument. Purposive schemes are again illustrative here. Against this reading of [1d], it is perfectly correct to say that Edmund's reason is to avoid thin ice in the middle of the pond. Here we do have a sentence — something that could be a part of the whole explanation — that adequately accounts for the phenomenon of Edmund's skating near the edge of the pond. But what this sentence gives as Edmund's reason (“to avoid thin ice in the middle of the pond”) is not itself a sentence, and it is not clear that it *could* be a sentence. So again, to the extent that we are impressed by the plausibility of purposive schemes, we should resist the view that either being a reason or being explanatory necessarily involves being a sentence.

This point about purposive explanations is suggestive of a more general reason to reject [1d*]. We might ask, “Why is Edmund skating near the edge of the pond?” as a way of pinpointing the phenomenon targeted by our explanatory efforts. Consider the following accounts of Edmund's skating near the edge of the pond.

“For the reason that the ice in the middle of the pond is thin.”

“To avoid thin ice in the middle of the pond.”

“In view of the danger posed by skating on thin ice.”

The second answer here is transparently purposive; but I want to focus on a different aspect of it here. All of these accounts are what I will refer to as **fragmentary**: they are not complete sentences (of natural language), but rather fragments. They also seem perfectly adequate as answers to the question posed. If you asked me why Edmund is doing what he is, my responding in one of these ways would not likely arouse any complaint or unclarity in your mind. But if the view put forth by [1d*] is correct, then we ought to find some deficiency in these responses *qua* explanatory accounts, and also as presentations of Edmund's reason.

We can sidestep the issue of what exactly is supposed to be wrong with these answers, since, on the view in question, the deficiency would presumably be remedied by transforming these accounts into proper sentences. As far as I can see, there are three sensible ways to do this. All three strategies rely on the suggestion that the fragmentary accounts are somehow shorthanded or elliptical; and that they in fact contain a hidden subject (the agent) which would be revealed if things were fully spelled out. One particular strategy is to suggest that the subject is somehow hiding in the operators that signal what the agent's reason is. So, according to this strategy, if we are to have the *real* explanatory claims, these fragments should be completed as follows.

“His reason is that the ice in the middle of the pond is thin.”

“His purpose is to avoid thin ice in the middle of the pond.”

“His consideration is the danger posed by skating on thin ice.”

We need not address the plausibility of the claim that the earlier fragmentary accounts are *really* shorthanded ways of expressing these sentential accounts. According to [1d*], a

reason is the kind of thing that can *be* a sentence in the whole *explanans*. So, in transforming our fragments into sentences, the product ought to be something that could be Edmund's reason for skating near the edge. But it is obvious that none of these sentences could *be* Edmund's reason. Edmund does not skate near the edge of the pond, for example, for the reason that *his reason is that* the ice in the middle of the pond is thin. Of course, this sentence might contain a clause representing or giving the content of Edmund's reason. But with that thought, we revert back to the "clausal" conception of [1d] we just discarded in a manner similar to [1a].

It seems that a similar problem arises with a second strategy, which operates on the suspicion that both the agent *and* his action are somehow hidden in the fuller meaning of the fragmentary accounts. Consider the following explanatory sentences.

"He is doing it for the reason that the ice in the middle of the pond is thin."

"He is doing it in order to avoid thin ice in the middle of the pond."

"He is doing it in view of the danger posed by skating on thin ice."

As with the sentences resulting from the first strategy, it is not plausible that any of these sentences are identical to Edmund's reason for skating near the edge of the pond, or even the content of Edmund's reason. It is not plausible, for example, that Edmund keeps to the edge for the reason that he is doing it in view of the danger posed by skating on thin ice. But this is just what the conception of the explanatory role of reasons expressed in [1d*] might lead us to if we insist that these fragmentary accounts are actually disguised sentential accounts of Edmund's reason.

There is a second, deeper problem with this strategy for [1d*] accommodating fragmentary explanations. It is clear that, even if the fragmentary accounts of Edmund's

skating are somehow deficient, there are satisfactory accounts very nearby. If [1d*] were true, we would expect that transforming these fragments into sentences would remedy this explanatory deficiency. However, it is difficult to locate any deficiency in the fragmentary explanations that could be remedied in the sentential accounts. The sentential accounts are not clearly better *qua* explanations of Edmund's action or presentations of his reason. Indeed, it seems clear that they are *not* any better. One hint that this is so is that the only revision is the addition of the clause "He is doing it;" but of course, we do not want to say that the fact that he is doing it can explain the fact that he is doing it. Neither should we like to say that "I am doing it" is any part of Edmund's reason for acting. So it seems that the clause added to transform these fragments into sentences does not contribute anything to the accounts *qua* explanations. Indeed, I struggle to see any explanatory task that is accomplished by the sentential accounts, but not by the fragmentary accounts. It is not as though, for example, the sentential explanations put us in a better position to understand what Edmund is doing, or give us a better grasp of what Edmund's reason is. But again: if [1d*] were true, the sentential accounts should have some clear explanatory advantage over the fragmentary accounts.

There is a third, more subtle strategy for transforming my fragmentary accounts into sentences. It relies on the thought, common among linguists, that infinitival clauses themselves contain a hidden subject.⁴⁰ Borrowing this thought, we could translate our second fragment as follows.⁴¹

⁴⁰ I am indebted to Ray Buchanan, Josh Dever, and Chris Hom here. See Liliane Haegeman [1991: 235ff] and Elizabeth Cowper [1992: 157ff] for standard accounts of this orthodoxy in government and binding theory. I am aware that there is a controversy about PRO theory in generative linguistics. But my impression is that even those linguists who think we can do without so-called "big PRO" (and get along fine with just so-called "little pro") hold that there is a hidden subject in infinitival clauses.

⁴¹ Since this linguistic account seems to hold only for infinitival clauses, it is not clear how it could be used to address the other two fragmentary explanations. For that reason I treat only the infinitival account. My impression is that even those linguists who think we can do without so-called "big PRO" (and get along fine with just so-called "little pro") hold that there is a hidden subject in infinitival clauses.

“[In order that] He avoids thin ice in the middle of the pond.”

The idea here would be that the proper answer to the question “Why is Edmund skating near the edge of the pond?” is given by the that-clause picked out by the phrase ‘in order that’. As before, we need not concern ourselves with whether this sentence is truly identical in content to the infinitival clause given our second fragment. Our questions remain the same: (i) Is it plausible that this sentence is either identical to or gives the content of Edmund’s reason for skating near the edge of the pond?; (ii) Does this sentential explanation have any advantages over the fragmentary explanation, such that we should insist upon it?; and (iii) Does this transformation into sentential form do anything to convince us that a motivating reason must be a fact in order to have explanatory power?

I see no principled reason not to allow a positive answer to question (i). Though one might think it odd that Edmund reference himself in giving his reason, there is no reason to think that with this addition we are no longer dealing with something that could be his reason for acting. In my view, Edmund could very well be motivated by the consideration that he will avoid thin ice in the middle of the pond. The answer to question (ii), though, remains negative, for the same reasons as mentioned previously. There is simply nothing to choose, vis-à-vis explanatory power, between the following two accounts of why Edmund is skating near the edge:

“In order to avoid thin ice in the middle of the pond;”

“[In order that] He avoids thin ice in the middle of the pond.”

It is not as though the mention of or reference to Edmund lends greater understanding to Edmund's action, or somehow improves the account such that it better captures the phenomenon of Edmund's skating near the edge (or Edmund's reason for so skating). But again, if [1d*] were true, we should expect to have gained an explanatory advantage by introducing the second account.

Finally, the answer to question (iii) also remains negative. There seems to me to be no more reason to think that "In order that" constructions are factive than there is to think that "In order to" constructions are factive. It is true that "In order that" constructions will be completed by propositions — things that *can* be true or false. On the other hand, "In order to" constructions are not completed by a clause that straightforwardly accepts a truth value. But the introduction of a that-clause by itself creates no factive pressure at all. We saw this when we considered psychologized explanations. The claim that "*A* Φ d in order that *p*" no more requires the truth of *p* than does the claim that "*A* Φ d because he believed that *p*." So, if we adopt [1d*] in a manner that allows the that-clauses in "In order that" contexts to count as explanatory sentences, then we show [2d*] to be implausible.

The points raised in this section, I think, give us good reason to believe that any variation of AFE that insists on any kind of sentential conception of motivating reasons, or their explanatory role, will fail. Whether one fills out this framework by saying that *being explanatory* is being the whole set of sentences comprising the *explanans*, being one of the sentences of that whole set, or being a part of one of those sentences, it seems that any version of AFE resting on this framework will include at least one implausible premise. So, as with our investigation of "Why?" and 'because', we simply do not find here the necessary resources for mounting a successful argument for the Factive View_M.

CONTRIBUTING TO AN EXPLANATION

So far, the considered accounts of what it is for a reason to be explanatory have relied heavily on analyzing things that can be thought of as pieces of language — questions, answers, sentences, and so on. But we can imagine a Factivist conceding, for instance, that *I* can explain Edmund’s action by *saying*, “For the reason that the ice in the middle of the pond is thin,” or “He is doing it in order to avoid skating on thin ice.” This can be conceded even while rejecting the idea that the false consideration that “the ice in the middle of the pond is thin” has been given a bonafide explanatory role. Here is Constantine Sandis:

... it is people who explain actions by *citing* one or more agential reasons, thereby implying strictly that (a) the agent took *p* and/or *q* to count in favour of her action and (b) acted accordingly. But the explanation is not done by the reason cited... Even in cases where we mistakenly agree with the agent’s [false] belief at the time of acting and therefore (successfully) explain her action with a simple ‘she did it because *p*’ what does the explaining is not the falsehood ‘that *p*’ but the implied truth that she acted upon the belief that it was. The explanation is only non-factive in the weak sense that a falsehood can feature in an explanatory statement that is true. The reason cited does not itself even contribute to the explanation. Just as the statement ‘this was her cat’ can, if true, explain why the animal followed her without the cat being an *explanans*, so the statement ‘this was her reason’ can explain why she did something, without the reason functioning as an *explanans*.⁴²

We can see here various rejections of some of the conceptions of the explanatory role of reasons we have treated so far. Clearly, Sandis thinks, conceptions such as [1a], [1b], [1c], and [1d] are out. What, then, is left? What seems to be left in Sandis’ treatment is the idea that what it would mean for a reason to play an explanatory role vis-à-vis an action is for the reason *itself* — not merely a statement of or including the reason — to *contribute* to the explanation of an action.

⁴² See Sandis [2013: 37].

With this thought, we arrive at the conception given by [1e], which leads us to yet another variation of AFE:

[1e] A reason is the kind of thing that can contribute to the explanation of an action.

[2e] Anything that can contribute to an explanation is a fact.

[3e] Therefore, all reasons are facts.

This variation of AFE is not vulnerable to many of the criticisms I have raised so far. There can be no worry, for example, about [2e] implying that explanatory claims or sentences cannot include false that-clauses; or about improper restrictions on the sort of thing that can be given in an answer to a certain kind of question; or about insisting that all explanations come in the same form.

The real force of this variation of AFE is that it concedes that reasons claims may not be factive, and may also be explanatory. But, while it may be true that we can explain Edmund's action, for example, with the claim "His reason for Φ ing is that the ice in the middle of the pond is thin," this should not lead us to think that ice's being thin is playing any explanatory role vis-à-vis his Φ ing. On the contrary, it is clear that in this explanation, the consideration that the ice is thin is not making an explanatory contribution, because it has no explanatory power of its own to contribute. We know that the consideration that "the ice in the middle of the pond is thin" has no explanatory power of its own because it could not by itself do anything to explain Edmund's action (because it is patently untrue, and only facts can explain things). But anything that is explanatory must itself be capable of at least contributing to an explanation. Since we

know that reasons are explanatory, we may then infer that anything that is not the case cannot be a reason.⁴³

Or so this variation of AFE would say. While I think this is, in a way, the most promising way to run AFE, I think that it ultimately falls short of its target. This is because [2e] does not seem to me to be true. To see why, we can revisit psychologized explanations of characters like Edmund. This present version of AFE makes no attempt to prevent the claim that “the ice in the middle of the pond is thin” from entering into the explanation of Edmund’s skating near the edge of the pond. Indeed, we have seen that no one should try to impose such a prohibition, because it seems that any good reasons-explanation of Edmund’s action must make use of the false that-clause in question. Recall that (most) Factivists simply opt to embed the false clause in an inoculating psychologized context (such as belief). Furthermore — and more importantly — it is clear that this psychological embedding does not (and is not meant to) make the content of the psychological state explanatorily irrelevant, because it is clear that not just any content can play the desired role here. The consideration about the thinness of the ice is unlike, for example, Edmund’s belief that George Washington was the first President of the United States. The latter is not poised to do anything to explain Edmund’s skating near the edge of the pond, but former is. What is under discussion here is, in a way, an example of what all action theorists are after in error cases: a way to allow untrue yet explanatorily relevant contents into explanations of actions without violating the axiom that explanation is factive.

⁴³ Sandis [2013] seems to arrive at a different conclusion, which is that these motivating reasons do not explain actions. This is perhaps of the same spirit as Alvarez [2010], who would like to make a distinction between motivating and explanatory reasons. Despite some of the complications noted here, I am somewhat more hesitant — in view of considerations of simplicity, perhaps — to move to a disjunctive account.

All of this, however, creates a puzzle for those who would use the version of AFE that has claim [1e] as a premise. Once we agree that the consideration that the ice is thin is, in this way, explanatorily relevant — indeed, it seems indispensable — vis-à-vis Edmund’s action, what could it amount to to claim that the consideration does not *contribute* anything to the explanation (or that it does not make an explanatory contribution)? What could it amount to to claim that some thing can (must?) be given in a good answer to the question “Why?”, be given in order to explain an action, be an indispensable part of any satisfying explanation, be used to explain an action, be a part of the *explanans*, and so on — and yet also claim that this thing does not contribute to the explanation of the action? It is not clear that there is here a sufficiently distinguished notion of *contributing to an explanation* that can be used to drive a wedge between the possible explanatory roles of true and false motivating considerations.

One thought might be that we can use a sort of *isolation test*, in the spirit of Moore’s test for whether some state or feature has intrinsic value.⁴⁴ In order for something to make a contribution to an explanation, it must have some explanatory power *of its own* to contribute; and we can tell whether something has explanatory power of its own by observing whether it can, on its own, serve to (at least partly) explain the phenomenon in question. Nothing that is not the case can explain anything on its own, because explanation is factive. So, false considerations, such as the one supposed to be Edmund’s reason for keeping near the edge of the pond, are shown to have no explanatory power of their own — and thus none to contribute to any explanation.

⁴⁴ See G. E. Moore [1903: 142, 145-157, 236, 256] and W. D. Ross [1930]. My line of resisting such tests is perhaps in the spirit of and Dancy [2004], and also McDowell [2013: 19], who observes that “it does not add to the explanatory power of explanations given by using those forms if things are as the agent relevantly takes them to be.”

This line of thought, though perhaps somewhat appealing, seems to me to miss most of the truth about how explanation works in general. Explanation, it seems to me, is *holistic*: the fact that some feature serves to explain something in one context may have little to no bearing on whether it serves to explain something (even that very same thing) in another context. This seems to me to be true even of ordinary causal explanation. The fact that it is cold outside does nothing, by itself, to explain the fire's being lit. It is only in the presence of other features — Frank's being home from work today, Frank's aim to keep the house warm, et cetera — that its being cold has any explanatory power vis-à-vis the fire's being lit. Now, together, these features provide the materials for a satisfying account as to why the fire is lit. There is no doubt about whether *its being cold* contributes to this explanatory account of the fire's being lit; and the fact that its being cold would not, by itself, do anything to explain the fire's being lit does nothing to affect our judgment that *its being cold* contributes to the more complex explanation.

So it seems to me that isolation tests of the sort considered here should hold little to no weight in determining whether any feature or consideration can contribute to any particular explanation. We are then left with the thoughts that false motivating considerations are essential to any good explanation of action in error cases, that they play a crucial role in helping us to understand the phenomena in question, and that they can lend rational intelligibility to an action even when cited in isolation. These thoughts are asked to face up to a second thought, that no thing that is not the case can *contribute* to an explanation. But we are left without an analysis of *contributing to an explanation* which should convince us that, despite the crucial role they play in action explanations, false motivating reasons do not contribute to action explanations. Admittedly, this has all proceeded without *my* supplying any analysis of what it means to contribute to an explanation, such that we can positively establish that false motivating reasons are

capable of doing so. My own view is that all motivating considerations — even false ones — play an explanatory role with respect to actions, and I will say just what that role comes to momentarily. But suffice it to say, for now, that there does not seem to be much sense in saying that motivating reasons play an explanatory role vis-à-vis actions, and then proceeding to deny that motivating reasons contribute to explanations of action. In light of the foregoing considerations, I suspect anyway that any analysis that leaves the notion of *contributing to an explanation* close enough to the intuitive notion will render [2e] implausible; but any analysis sufficiently tailored to guarantee [2e] will likely ring false when applied to reasons, rendering [1e] implausible. Either way, though, the burden is on the advocate of AFE to provide us an informative analysis which precludes false motivating considerations from playing an explanatory role. Unless one has been provided, the Nonfactivist can comfortably reject AFE by the reasoning laid out in the previous paragraphs. So, as in the previous two subsections, I think that advocates of AFE will need to look elsewhere to find the resources necessary for mounting a successful argument for the Factive View_M.

FACTUAL EXPLANATIONS FOR THE NONFACTIVE VIEW: A BRIEF SKETCH

I hope that at this point one can see a pattern emerging. We have examined at least five disambiguations of what it might mean for something to be explanatory. While to my knowledge these five disambiguations pretty well run the gamut of how things are discussed in the literature, it must of course be acknowledged that these disambiguations do not exhaust logical space; there may be other ways of disambiguating the notion of being explanatory, and therefore, there may be other ways of formulating AFE. But, as I said, one can see a pattern emerging. Advocates of AFE must disambiguate the notion of being explanatory. On pain of fallacy, they must proceed to revise both premises of their

general argument captured by [1]-[3]. But, once both premises are revised to include any satisfactory disambiguation, at least one of the two premises is seen to be implausible. While it is true that we have only examined a handful of options for the advocate of AFE, we have no reason to believe that any forthcoming disambiguation will deliver more promising results for the Factive View_M . We have at least reached a point, then, where we can reasonably require as much — the presentation of a more troubling disambiguation — of the Factivist adopting this strategy; and we can refuse his argument until he obliges.⁴⁵

Now, none of this shows that AFE is bound to fail (although it does perhaps suggest this, and I do suspect that it is bound to fail). But this all does make it difficult to see how the argument could succeed. It is difficult to see how adherents of the Factive View_M can succeed in articulating our two axioms (REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS) in a way that both retains their plausibility and puts them in a strict entailment relation to the Factive View_M . As I hinted at the beginning of this essay, though, I think that at the heart of the objection is a suspicion that the Nonfactive View_M is not compatible with satisfying reasons-explanations in all cases. Trying to show that the Nonfactive View_M is logically inconsistent with the two axioms in error cases is just one way of trying to formalize this deeper suspicion. While my main objective so far has been to demonstrate that AFE suffers from a fatal ambiguity, I want to conclude this chapter by going some of the way toward putting this suspicion to rest. I will do this by briefly sketching the kinds of satisfying reasons-explanations of actions that are available to the Nonfactivist, in accord with REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS.

⁴⁵ Those still feeling the pressure of this version of the argument may, like Sandis [2013: 37] reject one of the axioms in order to resolve the tension. But if something rightfully thought of as axiomatic is in doubt, and the argument thought to place it in doubt is also in doubt, it is better practice to keep the axiom and refuse the argument.

What is presented in this section will be only a sketch, to be completed once our inquiry in the third chapter gives us a fuller theory of explanation.

Conveniently, most of the explanations I want to focus on are just those which created trouble for AFE in the previous section. To get these explanations in view, though, we can reset to the beginning of the essay. Edmund is going skating, and he takes it that THIN ICE: the ice in the middle of the pond is thin. So, when he skates, he keeps to the edge of the pond. Unbeknownst to Edmund, the ice in the middle of the pond is quite thick and presents no actual danger. We might ask Edmund — inquiring about his reason for acting — why he is keeping to the edge of the pond. Edmund would probably say something like, “the ice in the middle of the pond is thin.” Suppose we, as theorists, now ask ourselves why Edmund is keeping to the edge of the pond. The Nonfactivist can use the materials provided by Edmund’s answer, and is uniquely poised to do so. Here are some explanations available to the Nonfactivist:

“He is doing it for the reason that THIN ICE.”

“He is doing it in order to avoid thin ice.”

“His reason is to avoid thin ice in the middle of the pond.”

“His purpose is to avoid falling through the ice in the middle of the pond.”

“He is aiming to avoid thin ice in the middle of the pond”

“He is skating so as to avoid dangerously thin ice.”

“He is doing it in view of the danger posed by thin ice.”

“He is responding to the consideration that THIN ICE.”

“He is doing it because his reason is that THIN ICE.”

“He is doing it so that he will not fall through the ice.”

For the sake of brevity, I will not give extensive treatment here as to just how I think each of these explanations is working (as I said, fuller treatment will come in the third chapter). But I do think that in each of these accounts we have satisfied any plausible conception of the axiom that reasons play an explanatory role — those suggested by [1a], [1a*], [1b*], [1d], and [1e], for example. These accounts explain Edmund’s action by specifying (*à la* Pryor) and making indispensable use of his reason for acting; we have explained his action by, one way or another, making clear and salient the consideration which motivated Edmund to act in the way he did. This is at least one significant sense in which the Nonfactive View_M respects the axiom of REASONS-EXPLANATIONS: it reserves a legitimate explanatory role for motivating reasons by holding them to be that which must be specified in any good reasons-explanation of an action done for reasons.

Furthermore, each of these answers seems to be in keeping with any plausible understanding of the axiom that explanation is factive, since each of these answers is true.⁴⁶ One might complain that these explanations, as I have put it, proceed by way of falsehood. But here we must recall once more that the explanations favored by Factivists in these cases also proceed by way of falsehood. For the Factivist’s explanation in this case will be something like:

“He is doing it because he believes that thin ice.”

According to (most) Factivists, this psychologized explanation is a good reasons-explanation, and we should say here that Edmund’s reason *just is* that he believes that

⁴⁶ Of course, it would be infelicitous for the critic of the Nonfactive View_M to object to this claim at this point (that is, without some other argument as to why falsehoods cannot be reasons).

THIN ICE.⁴⁷ But according to the Nonactive View_M I am sketching out here, Edmund's reason may be that THIN ICE; and we may present this as Edmund's reason by specifying *what* he believes without equating his reason with his belief. In the above explanations, the that-clause THIN ICE specifies the content of something crucial to a good reasons-explanation of what is doing. Due to the structural similarity between each of these explanations, I suspect that it will be difficult to generate any complaint about factivity which impugns the explanations just offered as available to the Nonfactivist, but not the psychologized explanations preferred by Factivists. If this remark seems to be *tu quoque*, it need not be. The point here is not that Factivists are companions in guilt with Nonfactivists. The point, rather, is that in error cases, every *good* reasons-explanation must proceed by way of falsehood at least in this weak sense. Since a good reasons-explanation will be one which helps us to see what the agent saw in doing what she did, the use of false clauses or considerations capturing what the agent saw is not a shared shortcoming, but a common strength, and the mark of a good explanation.

So, I would say that each of the above accounts satisfies REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS in a manner consistent with the truth of the Nonfactive View_M. Indeed, each of these accounts (save for the psychologized one) is available *only* to the Nonfactivist, since what they present to be Edmund's reason is not a fact. So the suspicion that the Nonfactive View_M is incompatible with good reasons-explanations in error cases is ultimately unfounded. Admittedly, this has only been a very brief sketch of the way in which the Nonfactivist could begin to fill out an explanatory program while respecting the two axioms we have been discussing; and a complete view of action, reasons, and explanation will say much more than this. The point of providing this sketch

⁴⁷ The qualification here is in observation of those Factivists like Alvarez [2010], and perhaps Alfred Mele [2007] and Sandis [2013], who opt to say of characters like Edmund that there is no (motivating) reason for which he acts (though there is a psychologized *explanatory* reason).

was simply to show that one going narrative in action theory is not true: AFE is quite far from taking the Nonfactive View_M off the table. Indeed, the Nonfactivist has good prospects of providing what AFE puts under scrutiny: good reasons-explanations of actions. In fact, there seems to be no reason to think that the Nonfactive View_M is *any worse off* than the Factive View_M as far as that goes. It may, of course, turn out that a Factivist program is in general more plausible than a Nonfactivist program in action theory; but this claim has certainly not been earned via AFE. Indeed, AFE has not even earned the Factivist an advantage vis-à-vis action explanation. What I take from all of this is that, all the dispute around AFE notwithstanding, the debate between Factivists and Nonfactivists will likely have to be settled along some other line.

Chapter Two: Acting in Unbelief

But how else might the debate be settled? For some time now, theory of acting for reasons has developed around three general questions. One question concerns the ontology of reasons: are reasons propositions, states of affairs, beliefs, desires, or something else? Relatedly: whatever else reasons are, are they causes? This present project comments on this question only negatively and somewhat indirectly, insofar as the first chapter offers arguments that motivating reasons need not be facts (and so therefore not *true* propositions, *obtaining* states of affairs, *occurrent* psychological states, or perhaps causes), and this chapter will terminate in an argument against the view that reasons are psychological states (such as beliefs and desires). A second question concerns the factivity of acting for a reason: must it be the case that r in order for A to Φ for the reason that r ? On this second question, I hope to have offered good grounds in the previous chapter for holding that acting for a reason is not factive. Finally, a third issue concerns the explanation of actions done for reasons: what must be true of the reasons for which agents act, such that they can play an explanatory role vis-à-vis the actions done in response to them? While this issue was addressed in the previous chapter, a fuller answer will come in the next chapter.

Quite recently, though, a fourth question has arisen — although, as I will suggest, it has not been sufficiently separated from the three questions just mentioned. This is the question of what I will call the *cognitive constraint* on acting for a reason. Currently, the discussion of this issue centers around the question of whether an agent must know some consideration to be true in order for that consideration to be a reason for which the agent acts. What I call the Knowledge View answers in the affirmative. But, as I will suggest, *that* question has been considered largely as a part of the question of whether acting for a

reason is factive (indeed, if we have good reason to accept the Knowledge View, then to that extent we should reject Nonfactive View_M I have so far defended). I think the issue of the cognitive constraint warrants thorough and independent discussion in its own right; and my first intention in this chapter is to try to initiate and set parameters for such a discussion. I will argue that, once we get a clear view of the question, the most plausible position is what I call the Belief View: in order for an agent to act for a reason, the agent need only believe that reason to be the case.

The course of the chapter is as follows. In the first section, I sift through some of the current literature and argue that there is an independent and significant question about the cognitive constraint on acting for a reason. I proceed to extract from the literature three general views on this question: the Knowledge View, the True Belief View, and the Belief View. I also recount some of the initial motivating arguments for each of these three views, and ask whether each of these views gives a legitimate answer to the question of the cognitive constraint. I observe that, so far, these views seem to have been motivated mostly by answering the other questions mentioned earlier, and perhaps should not even be counted as providing a distinct view on the cognitive constraint. So, in sections two, I devote some time thinking about the issue of the cognitive constraint independent of other issues in action theory. In section two, I argue that conceptual analysis of the notion of *acting for a reason* gives us some reason to think that the Belief View is correct, and no reason to think that any of its competitors is correct. I also argue that an initial analysis of cases is strongly suggestive of the Belief View: while its competitors seem to suffer from counterexamples, what *would* be a counterexample to the Belief View turns out to be an incoherent case.

The arguments from section two, I think, are enough to establish the Belief View as the intuitive view. But in sections three and four, I consider arguments that the

intuitive view is, alas, not the correct view. I first consider arguments that the Belief View is too weak. Specifically, I consider arguments that the factivity of acting for reasons and the nature of action explanation compel us toward a stronger cognitive constraint (such as the Knowledge View). I argue that even the best versions of these arguments should not compel us to give up the Belief View. In section four, I turn to arguments that even the Belief View is too strong. Although this possibility has not been given much consideration, I believe that there are examples that put pressure on even the Belief View. I argue that ultimately we have good theoretical reason not to weaken the cognitive constraint beyond belief, and that the apparently troubling examples can be explained away.

So, I conclude that, at least until further argument is presented, we should endorse the Belief View on the question of the cognitive constraint. In the fifth and final section, I end by asking whether the truth of the Belief View should give us cause to reconsider some views on other questions in action theory. I answer in the affirmative, and tend to one example: I argue that if the Belief View is true, then the view that reasons for acting are psychological states (often called Psychologism) is implausible.

CLEARING GROUND FOR THE QUESTION OF THE COGNITIVE CONSTRAINT

The central question of this chapter is what cognitive relation some agent A must stand in to some reason r in order for A to perform some action Φ for the reason that r . Those who provide a positive answer to this question will offer what I will call a **cognitive constraint on acting for a reason**. There is a circle of current discussion that hovers around this issue. It deals explicitly with the suggestion that in order for A to Φ for the reason that r , A must know that r . Jennifer Hornsby has said, for example, that “a condition of Φ -ing for the reason that p , when one believes that p , is that one *knows* that p .” This position has come to be known as the Knowledge View. Against this

thought, some, like Jonathan Dancy, endorse the Belief View, that “one can run for that reason [that the train is leaving] so long as one believes the train to be leaving.”⁴⁸

Why, then — given these explicit statements on what I am calling the cognitive constraint — do I imply that this central question has not been given thorough independent treatment, and characterize this debate as only *hovering* around the question? The answer is in the manner in which these interlocutors frame the issue of the cognitive constraint. Almost without exception, the issue of the cognitive constraint has been seen as a part of (or depending on) either the discussion about the ontology of motivating reasons, the explanatory nature of reasons, or else the factivity of acting for a reason. For example, Hornsby’s now well-known argument for the Knowledge View is really a part of a larger argument for a “disjunctive” conception of acting for reasons on which “there are cases and cases of acting for reasons; and not all of them conform to the conception that philosophers have of them when they think of actions as explained by beliefs and desires.” The well-known argument centers around just the case we spent the previous chapter examining. It concerns

Edmund who believes that the ice in the middle of the pond is dangerously thin, having been told so by a normally reliable friend. But Edmund’s friend didn’t want Edmund to skate in the middle of the pond (never mind why), so that he had told Edmund that the ice there was thin despite having no view about whether or not it actually was thin. Edmund, then, did not keep to the edge *because* the ice in the middle of the pond was thin... The fact that ice was thin does not explain Edmund’s acting, even though Edmund did believe that it was thin, and even though the fact that it was thin actually was a reason for him to stay at the edge.

This is supposed to show that a “reason has to have registered with an agent if they are to have acted on it. We see now what this amounts to. A condition of Φ -ing for the reason that p , when one believes that p , is that one *knows* that p .”

⁴⁸ See See Hornsby [2008: 251] and Dancy [2011: 345], respectively.

The structure of this argument is not immediately apparent.⁴⁹ But one thing that is apparent is that this argument for the Knowledge View depends almost entirely on — if not simply *amounts to* — an argument for a particular view of action explanation (on which *A*'s Φ ing cannot be explained by the reason that *r* unless *A* knows that *r*). This is apparent in that the heavy lifting in the argument is accomplished by the observation that a would-be candidate for Edmund's reason for acting is ruled out by its not being capable of figuring in a certain kind of 'because' claim, or of explaining Edmund's action. There seems to be a move here directly from the thought that Edmund does not know that the ice in the middle is thin, to the claim that Edmund does not keep to the edge *because* the ice is thin. This is strongly suggestive of the thought that what Edmund's ignorance of whether *p* precludes, in the first instance, is a certain way of explaining his action via *p*.

Those who give arguments against the Knowledge View also seem to weave the issue together with others. Here is Jonathan Dancy, summarizing one of his recent papers in which he argues against the Knowledge View.

This paper considers and rejects arguments that have been given in favor of the view that one can only act for the reason that *p* if one knows that *p*. This paper contrasts it with the view I hold, which is that one can act for the reason that *p* even if it is not the case that *p*.⁵⁰

As the summary suggests, what Dancy's project here depends on (and perhaps ultimately amounts to) showing that acting for a reason is not factive. The exercise of arguing against the Knowledge View is taken up as a part of that project. Ultimately, Dancy seems primarily concerned (like Hornsby) to argue for a certain framework for explaining

⁴⁹ See Hornsby [2008: 244, 251] for the passages here.

⁵⁰ See Dancy [2011: 345]

action. Indeed, he proceeds: “I will begin by elaborating the two fringe positions, starting with my own. There is more than one way of explaining an action...”

For my part, I find this way of setting things up to be puzzling; and I invite you to be puzzle with me. Here is Dancy again, laying out a bit more of the terrain.

This Discussion of the relation between knowledge and acting for a reason has become much more complicated recently. Here are some views which it is well worth keeping apart:

1. You can act for the reason that p without knowing that p , but if you do so you are in breach of a norm.
2. You cannot act for the reason that p if you do not know that p .
3. You can only act for the reason that p if it is the case that p . Otherwise your reason will be something else, probably that you believe that p .
4. You can act for the reason that p whether it is the case that p or not, so long as you believe that p .
5. You can only be guided by the fact that p if you know that p . Otherwise you are merely treating p as a premise, and to do that you do not need to know or even believe that p .⁵¹

These five views are included in a discussion ostensibly concerning what I have called the cognitive constraint on acting for a reason. Dancy characterizes them explicitly as being positions in the “discussion of the relation between knowledge and acting for a reason.” But only one of the five views mentioned here — the second view — seems to be directly and genuinely engaging with the question of the cognitive constraint. Indeed, it is not clear that there is any single question on which all five of these views provide a clear answer. The first view is really concerned to argue for some *normative* constraint on acting for a reason. The third view seems not to even approach the question of the cognitive constraint; it instead discusses the factivity of acting for a reason, as well as the ontology of

⁵¹ See Dancy [2014: 81].

reasons. The fourth view is the view discussed in the previous quote, which advocates the Belief View only incidentally. Finally, the fifth view does not transparently discuss acting for a reason (though it does discuss two other phenomena which may be similar to acting for a reason).

What is the Issue of the Cognitive Constraint?

I take the manner in which Hornsby and Dancy proceed to be representative of the contemporary debate concerning the cognitive constraint. Now, allow me to cancel any implicature that these philosophers are deeply misguided in their arguments, or that they are mistakenly connecting issues that are obviously unrelated:⁵² the issue of the cognitive constraint clearly *is* related to these other issues. At the very least, they all concern the nature of acting for a reason. But it is also an interesting and important subject in its own right; and what I am suggesting here is that it has not been treated as such. I am also suggesting that philosophers sometimes unwarily alternate between these issues; and I think that this incautiousness sometimes leads to serious mistakes. I want to suggest, in the interest of avoiding such mistakes, that there are a handful of questions which, despite their being related (some more intimately than others), are well worth keeping apart:

1. What cognitive relation must A stand in to r in order for A to Φ for the reason that r ?
2. What cognitive relation must A stand in to r for r to play a role in explaining A 's Φ ing?

⁵² Allow me also to cancel any implicature that Dancy is somehow responsible for these convolutions. His original target, I take it, was what he called Psychologism about reasons. Dancy [2000] argued against this view in part by arguing that a motivating reason was not a belief that something is the case, but instead, a thing believed to be the case. Acting for the reason that r , on this view, obviously requires believing that r .

3. Can $A \Phi$ for the reason that r if it is not the case that r ?
4. Can the reason that r play a role in explaining A 's Φ ing if it is not the case that r ?
5. Can $A \Phi$ for the reason that r if r does not favor A 's Φ ing?
6. Can r be a reason for (that is, count in favor of) Φ ing if it is not the case that r ?
7. Can r play a role in explaining A 's Φ ing if r does not favor A 's Φ ?

It is the first question here that is of central interest to this chapter. This question is clearly important in its own right. But, as one can see in the passages above, this question is often addressed as a byproduct of answering others.

As I said, all of these questions, though distinct, are related. So, despite the unfortunate state of the current discussion, theorists *have* managed to offer views that in some sense should count as answers to the first question. With some work, I think we can extract from Dancy's list what are presently the three main competitors on the question of the cognitive constraint. The first, which is most prominently defended by Dancy himself, is the

Belief View: in order for A to Φ for the reason that r , A must believe that r .⁵³

Of the three main views, the Belief View advocates the weakest cognitive constraint. A second, stronger cognitive constraint is the

True Belief View: in order for A to Φ for the reason that r , A must believe that r , and it must be the case that r .

⁵³ See Dancy [2000; 2011; 2014]. For a relevant endorsement of the Belief View for reasons for belief in particular, see David Enoch [2010: 981ff].

I think it is safe to say that this second view is held by the majority of theorists who have considered the nature of acting for a reason. This is perhaps partly in virtue of the Factive View_M being in the clear majority. For there is another kind of Factive View_M, called the

Knowledge View: in order for A to Φ for the reason that r , A must know that r .

The majority of those who endorse the Factive View, it seems to me, reject the Knowledge View in favor of something like the True Belief View.⁵⁴ The Knowledge View, like the Belief View, is a minority view — though it is not so far in the minority as the Belief View. Nonetheless, it has several prominent defenders, including Jennifer Hornsby, John Hyman, John McDowell, Peter Unger, and Timothy Williamson.⁵⁵

It is worth clarifying that none of these views purports to be telling the whole story about how an agent comes to act for a reason. Each of these views only claims that the cognitive constraint in question places a necessary condition on acting for a reason. As such, the Knowledge View will entail both the True Belief View and the Belief View (according to the orthodoxy which tells us that “ A knows that p ” entails both “ A believes that p ” and “ P ”); and the True Belief View will entail the Belief View.

⁵⁴ At least initially, we can put those who subscribe to a picture roughly following Davidson in this camp; but as we will see, those who hold that motivating reasons are psychological states will ultimately have to reject any cognitive constraint on acting for a reason.

⁵⁵ See Hornsby [2008], Hyman [1999; 2011; 2015], McDowell [2013], Unger [1975], and Williamson [201X]. Littlejohn should perhaps be mentioned, along with, according to Dancy [2011: 346-348], H. A. Prichard [2002]. Alvarez [2010: 132fn9] seems prepared to admit that Hyman’s view might be correct, and thinks that her view could accommodate that possibility.

What's Truth Got to Do With It?

Despite its being the mainstream view on the issue, it strikes me as odd to take the True Belief View as a view about the cognitive constraint. The constraint advocated by the True Belief View has two aspects. The first is A 's believing that r ; the second is its being the case that r . This second constraint is already represented in this essay as an independent view of its own. It is the

Factive View_M: in order for A to Φ for the reason that r , it must be the case that r .

The oddity of the True Belief View *qua* position on the cognitive constraint is in this second aspect of the constraint. The oddity is that, simply put, truth is not a cognitive relation. In fact, it is not immediately clear what the Factive View_M has to do with the question of the cognitive constraint. The above list can be of use here. Recall that the question of interest is just: What cognitive relation must A stand in to r in order for A to Φ for the reason that r ? We should be quite puzzled by anyone who responded to this question by saying, "Well, it must be the case that r ." Indeed, one who responds in this way seems clearly to be providing an answer to the *third* question above — the question of whether acting for a reason is factive. The Factive View_M answers *this* question in the affirmative, but seems not to provide any answer to the question of the cognitive constraint. What is more, the Factive View_M does not seem to even *imply* a particular answer to the question of the cognitive constraint. Indeed, the Factive View_M seems to me consistent with every plausible view on the issue. Obviously, the Factive View_M is inconsistent with the view that, for example, in order for A to Φ for the reason that r , A must falsely believe that r . But no one even entertains such a view. If we limit our focus to the three main views of the cognitive constraint just mentioned, it is clear that the Factive

View_M does not imply the truth of any of them. We can see this easily: for any r , its being the case that r simply does not entail that A believes that r , that A truly believes (that is, has a true belief) that r , or that A knows that r .

It is true that the truth of the Knowledge View entails the truth of the Factive View_M (again, under the orthodox assumption that knowledge is factive). This might lead one to think that, if the Factive View_M is true, then the Knowledge View enjoys the advantage of being able to explain this fact in a way that neither the Belief View nor the True Belief View can. Furthermore, if, as I argued in the previous chapter, the Nonfactive View_M is correct, then neither the True Belief View nor the Knowledge View can be correct. I will address this thought in the third section here. For now, it is enough to observe that the Factive View_M itself provides no answer, either directly or indirectly, to the question of the cognitive constraint. This observation is more dialectically significant than it might first appear. For we are now in a position to see clearly that True Belief View is really a packaging of two different views on two different questions. One of these is the Factive View_M, which constrains acting for a reason by saying that the reason must be something that is the case. But, as we have just seen, this constraint neither is nor implies any *cognitive* constraint on acting for a reason; and we saw in the previous chapter that it can be given extensive treatment independent of any consideration of the cognitive constraint. Now, the other aspect of the True Belief View — A 's believing that r — *does* place a genuinely cognitive constraint on acting for a reason. But that constraint is just identical to the Belief View: it says that in order for A to Φ for the reason that r , A must believe that r .

So the True Belief View is just the conjunction of the Factive View_M and the Belief View. The Factive View_M provides no answer to the question of the cognitive constraint, while the Belief View does. What all of this shows, I think, is that, so far as the

question of the cognitive constraint is concerned, the True Belief View just collapses into the Belief View. So there are really only two genuine competitors on this issue: the Belief View and the Knowledge View. Somewhat surprisingly then, what is often considered the mainstream view on the cognitive constraint should actually be excluded from the discussion — and that is more or less what I will do for the remainder of this chapter. Before beginning my defense of the Belief View, though, I want to address one tempting thought about the dialectic here. Given what I have just said about the True Belief View, one might be tempted to also declare that the Knowledge View is also not genuinely a view on the cognitive constraint of acting for a reason. On what grounds? The suspicion might be that knowledge is not *purely* a cognitive state, as belief is. Recall that our question is: what must be going on cognitively in A such that A can Φ for the reason that r ? Now suppose for a moment that knowledge is just justified true belief. On this conception, the Knowledge View can be seen as having three aspects, just as the True Belief View can be seen as having two aspects. One of these aspects — truth — is neither here nor there with respect to the cognitive constraint; the truth of the Factive View_M does not imply, for example, that the Knowledge View is true and Belief View is false. Channeling Putnam, we can say: the cognitive constraint is about what is going on inside an agent's head, and truth just ain't in the head.⁵⁶

Putting the factive aspect of the Knowledge View aside, then, we can consider the other two aspects of the view. These aspects might be captured by what we can call the

Justified Belief View: in order for A to Φ for the reason that r , A must believe that r , and A must have justification for believing that r .

⁵⁶ Hilary Putnam [1973: 704], of course, made his memorable remark about *meaning*.

The Justified Belief View, like the True Belief View, appears to be a package of two views. One view is the Belief View, which I think everyone will acknowledge as a purely cognitive constraint on acting for a reason — belief surely is in the head. The other is the

Justification View: in order for A to Φ for the reason that r , A must have justification for believing (or be justified in believing) that r .

The Justification View, as with the Factive View_M, might seem to be silent on the question of the cognitive constraint, at least in the following way. Supposing that we ask a theorist whether A is Φ ing for the (motivating) reason that r , we would be quite puzzled if the theorist pursued this inquiry thusly: “Well, let us see. Does A have justification for believing that r ?” At least initially, we might think this theorist has changed the subject. It is of course possible to hold that if A is not justified in believing that r , then A cannot Φ for the reason that r without breaching some *norm* of action. This seems to be the position of John Hawthorne and Jason Stanley, for instance.⁵⁷ But in considering this thought we shift away from the issue of the cognitive constraint toward another issue (perhaps that identified by the first view on Dancy’s earlier list). Furthermore, it is clear that one can bear a number of significant cognitive relations to r (including belief) without having justification for standing in that relation to r . We can generalize here, and observe that there are a number of significant cognitive states (belief included) that simply contain no *normative* element whatsoever. Finally, it is not clear that *having justification* is itself a cognitive state. So we might be inclined to think that justification, like truth, just ain’t in

⁵⁷ See Hawthorne and Stanley [2008].

the head — and so it cannot possibly be a necessary element of any *cognitive* constraint on acting for a reason.

The Justification View, then, does not obviously provide a cognitive constraint on acting for a reason, and it is separable from the only aspect of the Knowledge View that *does* clearly provide such a constraint (the Belief View). Does this not create a predicament for the Knowledge View similar to that of the True Belief View? After all, we are now in a position to say that the Knowledge View is really a package of three views — the Factive View_M, the Belief View, and the Justification View — each of which seem to place three different constraints on acting for a reason. But neither the Factive View_M nor the Justification View seem to provide *cognitive* constraints on acting for a reason. Should we not then say, just as we did with the True Belief View, that so far as the question of the cognitive constraint is concerned, the Knowledge View really just collapses into the Belief View?

I do not think that things are quite that simple. I offer two softening considerations that I believe should allow the Knowledge View to stay in the game, as it were. The first softening consideration I offer is that this predicament for the Knowledge View, as I have raised it, might rely on adopting a certain view of epistemic justification. The Factive View_M seems uncontroversially separable from the Belief View in a way that the Justification View is not. Whether my belief is true seems to be answerable to something over and above what is going on in me cognitively. But it might not be so for whether my belief is justified. On certain views of epistemic justification, whether my belief is justified might be answerable only to other cognitive facts about me; I need no cooperation from the external world in order to have a justified belief. Justification *might be* just in the head. If this is true, then it is not clear that the Justified Belief View is not a genuine answer to the question of the cognitive constraint.

The second softening consideration I offer is that this predicament for the Knowledge View, as I have raised it, relies on analyzing knowledge as justified true belief. But this once-orthodox view has fallen somewhat out of favor, with some going so far as to deny that knowledge has *any* true analysis. Perhaps it is true, as Williamson says, that Knowledge is *merely* a state of mind. This claim may be unexpected. On the standard view, believing is merely a state of mind but knowing is not, because it is factive: truth is a non-mental component of knowing.

Factive mental states are important to us as states whose essence includes a matching between mind and world, and knowing is important to us as the most general factive stative attitude.

The factive mental states are as genuinely mental as any states are.⁵⁸

On this view, the factivity (and perhaps the justification) of knowledge is *constitutive* of the kind of mental state that it is. If this conception of knowledge is correct, then it would be incorrect to characterize the Knowledge View as merely cobbling together the Factive View_M, Belief View, and Justification View. Rather than being a kind of belief, knowledge might be a genuinely cognitive state in its own right, completely distinct from belief. This can be true, of course, even if “*A* knows that *p*” entails that “*A* believes that *p*”. A similar remark applies to justification. Even those who deny that there is any true analysis of knowledge will be quite tempted to hold that claims of the form “*A* knows that *p*” entail claims of the form “*A* has justification for believing that *p*.” So the truth of the Knowledge View will rest on the truth of both the Justification View and the Belief View. But this does not show that knowledge is a kind of justified belief. For it might be that justification and belief are partly *constitutive* of knowledge, which is a mental state, genuinely cognitive in its own right.

⁵⁸ See Williamson [2000: 20-21, 39, 53] for the three passages here.

In any case, the insight here is that there is a legitimate conception of knowledge on which the strategy of reducing the Knowledge View, vis-à-vis the cognitive constraint, to the Belief View would be hopeless. It is difficult to find similar grounds on which to resist the thought that the True Belief View is *just* a cobbling together of the Factive View_M and the Belief View. In short, we should not like our treatment of the cognitive constraint to prejudge a controversial issue in the metaphysics of epistemology. In that spirit, I think that we should count the Knowledge View as offering a genuinely cognitive constraint on acting for a reason.

Even setting these softening considerations aside, there is anyway an important question about whether the Knowledge View is true, and whether there is any argument to advantage it over and above the Belief View. Furthermore, there is a strong current of thought in the literature that the Belief View and the Knowledge View are opposed on *some* philosophical question, even if it is not exactly the one I have identified here. We can certainly wade into that debate while registering our doubt that there is any single question on which the two views are directly opposed.⁵⁹ With all of this in mind, I think that, while one might reasonably suspect that the Knowledge View is something of an imposter in the conversation on the cognitive constraint, it would be unreasonable to exclude it from the debate in the same way that I have excluded the True Belief View.

⁵⁹ I have not addressed the suggestion that the two views might be opposed on the truth of the Justification View. But the dialectical situation there is the same as it is with respect to the Factive View_M. The Knowledge View entails the Justification View; but the truth of the Justification View does not entail the falsity of the Belief View. Indeed, the possibility of the Justified Belief View demonstrates this. So, while the Knowledge View insists on a normative constraint on acting for a reason, the Belief View is silent on such constraints. It seems to me much too strong to take this as grounds for declaring the two views to be directly opposed on the general question of the cognitive constraint.

AN INITIAL DEFENSE OF THE BELIEF VIEW

So far I have tried to clear some ground around the issue of the cognitive constraint on acting for a reason, so that we can both see it as a significant and distinct issue, and also get a clear view of it. As I will further explain in this section, the issue homes in on a particular question about the nature of *motivation*: what cognitive relation must an agent bear to some consideration r in order to be motivated by that consideration to act? I have suggested that discussion of this question has often been, perhaps unwarily, subsumed under discussions about the factivity of acting for a reason and the nature of action explanation. An unfortunate result of this is that there may be some views in the literature (such as the True Belief View) which are merely masquerading as views on the cognitive constraint. These initial considerations have lead me to believe that the real debate over the cognitive constraint is one between the Belief View and the Knowledge View. In this section I will try to lend greater clarity to this debate by approaching the issue more directly. I begin by giving an initial and novel defense of the Belief View. My strategy is quite simple: I argue that both conceptual analysis of *acting for a reason* and a simple examination of cases establishes the Belief View as the intuitively correct view of the cognitive constraint.

Acting for a Reason: The Very Idea

Just as it was in considering the question of whether acting for a reason is factive, in looking for a cognitive constraint on acting for a reason it seems sensible to devote some time to examining the notion of acting for a reason. Here we can make use of some of the work done in the previous chapter. What I aim to show here is that, of the standard analyses of acting for a reason, none are suggestive of the Knowledge View, and some are suggestive of the Belief View.

I would like to start with the notion of responding to a consideration (where, as before, a “consideration” is understood as a thing considered). Acting for a reason is, *inter alia*, a way of responding to a consideration. What can we say of cognitive constraints on responding to a consideration? Now, some might find it immediately transparent that one cannot respond to anything that one does not believe to be the case, and also that one can respond to a consideration that one does not know to be the case. But a look at an example might shed some light on the issue for those to whom this is not so transparent. We can return to the example involving your friend Aria, who has called off her wedding. You tell me that Aria has called off the wedding, and I ask why. You might tell me that she has done it in view of (or in response to) her fiancé’s unfaithfulness, or alternatively, that she is responding to (or doing it in view of) the consideration that her fiancé has been unfaithful. Now it seems to me obvious that in saying any of these things, you are not saying or implying that Aria *knows* that her fiancé has been unfaithful. What you have told me entails absolutely nothing about what Aria does or does not know. For those in doubt, consider how natural and unproblematic it would be for me to respond to you by asking, “Well, does she know that he was unfaithful? How does she know?” If knowing that *r* was somehow conceptually connected to responding to the consideration that *r*, we would expect such questions to at least strike us as infelicitous. But they clearly do not strike us so. Saying that Aria is responding to the consideration that her fiancé has been unfaithful does, however, seem to presume or imply that Aria *believes* that her fiancé has been unfaithful. While I could unproblematically ask you whether Aria knows that her fiancé has been unfaithful, it would be nearly incoherent for you to continue your answer to me by saying, “No; she does not know. In fact, she does not even believe that her fiancé has been unfaithful. Nonetheless, that consideration has come to her attention, and this is how she is responding to it.”

So I think it is clear that while there might be a belief constraint on *responding to a consideration*, there does not seem to be any knowledge constraint on the phenomenon. It is also clear that acting for a reason is a way of responding to a consideration. Though I think this is suggestive, it does not strictly follow that the Belief View is true and the Knowledge View is false. That is because although acting for a reason seems clearly to be *a way* of responding to a consideration, acting for a reason and responding to a consideration are not identical: acting for a reason might be something more than just responding to a consideration. Consider, for example, that one can respond to a consideration by ignoring or rejecting it; but ignoring and rejecting some thing seem to preclude treating that thing as a reason (and perhaps also believing that thing to be the case).

What could we add to the notion of *responding to a consideration* to get what we think of as *acting for a reason*? A natural strategy here would be to specify either the kind of response or the kind of consideration that is peculiar to acting for a reason. We saw in the previous chapter that a common conception concerning the latter is that when A Φ s for the reason that r , r is a consideration A takes to favor A 's Φ ing. So we might think that, whatever else acting for a reason is, it is *responding to a consideration taken to favor* the action in question. Unfortunately, I think that adding this qualification to our analysis does not do much to clarify the debate concerning the cognitive constraint. Take the consideration that rain: it is raining. Now, I take it that we all think, *ceteris paribus*, that the consideration that rain favors bringing an umbrella. But surely one need not know that rain in order to either take it to be the case rain, or to take it that rain favors bringing an umbrella. Neither does it seem that one needs to believe that rain in order to take it that rain favors bringing an umbrella. Indeed, it seems to me that one could even believe that it is not the

case that rain while still taking it that rain favors bringing an umbrella.⁶⁰ I know of no way to prove this conclusively, but I do think an analogy to logical implication is instructive here. Placing considerations in favoring relations seems to me to be in some ways similar placing them in implication relations. One can take it that p entails q without knowing or believing that p — indeed, one can do this while believing or knowing that *not* p . It seems to me that the ability to take one consideration as favoring another is similarly insensitive to these cognitive facts.

Now, if we are not careful here, these considerations might seem to mount the beginning of an argument against the Belief View. This provides an opportunity to say more about what I take *responding to a consideration* to amount to. Earlier I suggested that one might respond to the consideration that r by rejecting or not believing that r ; this may seem to show that responding to a consideration does not entail believing that consideration to be the case. But that would be to move a bit too hastily. For our notion of responding to a consideration to approximate the notion of acting for a reason, we need it to be such that the consideration *itself* motivates the response in question. Now, what could motivate us to reject, for example, the consideration that rain? It would be quite a confusing character who was motivated by the consideration that rain *itself* to believe that it is not the case that rain. Instead, someone who rejects this consideration is probably motivated by, for example, how things look outside, what the weather report said, or some other such. This seems to me to be so even for considerations that are necessarily, self-evidently, or analytically false. What considerations motivate me to reject the consideration that two and two make five? Perhaps it is more tempting here to say that the motivation comes from the consideration *itself*: I consider the proposition, and

⁶⁰ Perhaps some want to say that when one believes that it is not the case that rain, one takes it that *if it were the case* that rain, then the consideration rain that would favor bringing an umbrella. I address this difference in the fourth chapter.

nothing else is required for me to be moved to reject it. But I think it is more plausible to say that we are motivated to reject such considerations in view of their incoherence, severe implausibility, or seeming incorrectness. Such features of these considerations (or the consideration that these considerations have such features or properties) are of course distinct from the considerations themselves. It is not as though, for example, *its being necessarily false that two and two make five just is* the consideration that two and two make five. Furthermore, it remains in such cases that I need not *know*, for example, that it is necessarily false that two and two make five in order to reject that two and two make five in response to the consideration that it is necessarily false. At most, I need believe it to be necessarily false. So I think that the notion of responding to a consideration actually provides us resources to more ambitiously say that the truth of the Belief View is suggested here, while the Knowledge View is positively *not* suggested.

The earlier analogy between logical implication and the favoring relation perhaps brings us closer to another conception of acting for a reason which we considered in the first chapter. This is the conception of treating something as a premise. We saw earlier, as Hyman rightly remarked, that it need not be the case that r in order for A to treat r as a premise in practical reasoning or deliberation. Is there any connection between this notion and the issue of the cognitive constraint on acting for a reason? Here again is Hyman, invoking at once the notion of treating something as a premise and the notion of responding to a consideration taken to favor a response:

A consideration that makes an action appropriate, from the agent's point of view, need not be true. But it need not be something the agent believes to be true either. It is helpful to call it a "premise" because it is clear that we can act on premises we do not believe. For example, I can act on the premise that a student felt too ill to come to a lecture—in other words, I can assume or accept that the student felt too ill—even if I do not believe this to be true.⁶¹

⁶¹ See Hyman [2011: 360].

I think Hyman is right that if acting for a reason is *merely* treating a consideration as a premise in reasoning, then the Belief View (and, *a fortiori*, the Knowledge View) seems implausible. Suppose I am wondering whether to carry an umbrella today. I might reason as follows: if rain, then I should carry an umbrella. Clearly, one can reason in this way without knowing or even believing that rain. It seems true, then, that the conception of acting for a reason as treating something as a premise in reasoning advantages neither the Belief View nor the Knowledge View. In my view, though, neither does this conception give us reason to doubt either the Belief View or the Knowledge View. Not every bit of reasoning done prior to acting straightforwardly reveals something about what one's reason for acting — the consideration that motivated one to act — is. For example, I might reason that I should carry an umbrella if it is the case that rain, but then look outside and notice that it is not the case that rain. I might then proceed to carry an umbrella to give me shelter from the sun. It would be absurd in this case to suggest that I am carrying an umbrella for the reason that rain, even if it would be sensible to say that rain was a consideration I treated as a premise in my deliberation. Hyman's insight about cognitive constraints (or lack thereof) on treating something as a premise, then, do not seem to extend to the subject of acting for a reason. Indeed, he seems to recognize this, as he proceeds to argue for the Knowledge View despite his remarks about treating a consideration as a premise.

Even though we have only examined a few analyses of acting for a reason, I think we have done enough to spot a trend. Such analyses will sometimes reveal belief that r to be necessary, and will never reveal knowledge that r to be necessary, for acting for the reason that r . At the very worst for the Belief View, such analyses may come out neither helping nor hurting either view. We could examine other analyses of acting for a reason, such as *being motivated by a consideration*, or even *doing that action taken to be favored by a*

consideration. But I think that they will all continue this trend. Conceptual analysis of acting for a reason provides some reason to hold that the Belief View is correct, and that the Knowledge View is incorrect.

Two Limiting Cases

So I think an initial analysis of the notion of acting for a reason leads us to think that the Belief View is more plausible than the Knowledge View. I want in this section to solidify the intuitive case for the Belief View by looking at a few examples. I think that a quick examination of cases reveals what seem to be clear counterexamples to the Knowledge View, and no clear counterexamples to the Belief View. Indeed, what *would* be a counterexample to the Belief View looks to be an impossible case.

Before looking at what I want to hold up as two limiting cases, I want to issue a reminder of what happened in the last section. Despite my insistence that we should try to set aside the question of factivity when thinking about the cognitive constraint, there are some connections between the two which cannot be ignored. First, the truth of the

Knowledge View: in order for A to Φ for the reason that r , A must know that r .
entails the truth of the

Factive View_M: in order for A to Φ for the reason that r , it must be the case that r .

Second, on the other side of the same coin, the truth of the Nonfactive View_M entails the falsity of the Knowledge View. I mention this because we saw in the previous chapter that the initial conceptual analysis and casuistry is heavily suggestive of the Nonfactive View_M.

So, in addition to any intuitive burdens incurred by the Knowledge View *qua* view on the cognitive constraint, it seems the Knowledge View incurs intuitive burdens *qua* view on the factivity of acting for a reason.

That brief interjection aside, here is a useful case for thinking about the Knowledge View independent of the Factive View_M.

RACIST JUDGE: Judge Jerry is deliberating about whether Gil is guilty of a crime. Gil did indeed commit the crime in question. Jerry arrives at the verdict that Gil is guilty. Although he is quite confident — nearly certain — in his verdict, his having arrived at this verdict is ultimately owed to an unjustified prejudice against people of Gil's race. This prejudice influenced Jerry to evaluate the evidence as indicating Gil's guilt, even though the evidence did not in fact indicate this. Jerry is now tasked with sentencing Gil, and he takes it that the appropriate sentence for the crime in question is two years in prison. So, Jerry sentences Gil to two years in prison.

Now, even without doing any sophisticated theorizing, I think that it is quite natural to describe this case as one wherein Jerry sentences Gil to two years in prison at least partly *for the reason* that HE DID IT!: Gil committed a crime the legally appropriate sentence for which is two years in prison. In case there is any doubt about this intuitive thought, though, we can call upon our previous analyses of acting for a reason. It also seems correct to say, for example, that Jerry sentences Gil to two years *in light of the consideration that* HE DID IT!; that in giving Gil two years, Jerry is responding to the consideration that he did it!; that Jerry takes the consideration that HE DID IT! to favor sentencing Gil to two years; that in deliberating about what sentence to hand down, Jerry treats the

consideration that HE DID IT! as a premise; and that Gil is motivated by the consideration that HE DID IT! to sentence Gil to two years.

So, we can say that Jerry acts for the reason that HE DID IT!. Jerry's belief that HE DID IT! is clearly unjustified, though, since it is the result of racist prejudices. So Jerry does not *know* that HE DID IT!, since, if he did know that HE DID IT!, his belief that HE DID IT! would be justified. However, this lack of justification seems not to have any practical importance. Indeed, we can imagine a reporter questioning Jerry about his perceived harsh sentence and asking Jerry *why* he sentenced Gil to two years in prison. He would likely respond with some appeal to the consideration that HE DID IT!. Expressing some doubt, the reporter might ask Jerry, "But how do you know that Gil is guilty?" to which Jerry might respond, "Well, I may not *know* that he is guilty; but that is what the evidence leads me to believe — and anyway that is why I handed down the sentence that I did." I take it that we would all accept this response from Jerry without much fuss. The response makes sense, even to those of us who know all the details of the case (including Jerry's racism); no principles of charity or interpretative gymnastics are required here. Jerry is speaking plain English, and we all understand him.

Suppose now that we remove from the case that Jerry's belief that HE DID IT! is the result of unjustified prejudices (but is instead the result of, for example, Jerry's training in jurisprudence and legally evaluating evidence). However, suppose that two eyewitnesses to the crime were crucial to Jerry's finding the case the way he did. Now, though these eyewitnesses are extremely reliable (they have never been wrong before, we can suppose), they lied on the stand (never mind why) in saying that Gil committed the crime. None of this makes any difference. It is not as though our judgment that Jerry acts for the reason that HE DID IT! is now either strengthened or weakened. These changes do not, for example, affect our ability see how Jerry could have been motivated by the consideration

that HE DID IT!, or our ability to understand his sentencing by appeal to the consideration that HE DID IT!. Whether present or absent, the justification of Jerry's belief that HE DID IT! seems to be neither here nor there when considering whether Jerry acted for the reason that HE DID IT!.

But all of this is at odds with the Knowledge View. If the Knowledge View is correct, then Jerry could not have acted for the reason that HE DID IT!. So, if the Knowledge View is correct, we certainly cannot accept Jerry's report that he sentenced Gil for the reason that HE DID IT!. Even more striking, though, is that if the Knowledge View is correct, then " $A \Phi$ s for the reason that r " entails or implies " A knows that r ." If that were the case, then Jerry's response to the reporter would not just be false or inaccurate — it is also incoherent, and should strike us as unintelligible, since we know that Jerry does not know that HE DID IT!. But as we just saw, Jerry's response seems sensible, and even correct. It certainly does not even approach anything like unintelligibility. Why, then, do we see remarks to the contrary in the literature? According to Peter Unger, for example, "It is inconsistent to say 'Ralph's *reason* was that Fred's hat was wet, but he *wasn't* absolutely *certain* that it was."⁶²

We can set aside the idiosyncratic remark about certainty, and suppose only that Unger had made his remark about knowledge. I think that we can treat such claims in just the same way that we treated in the previous chapter claims that it is paradoxical to say, " A 's reason was that r , but it is not the case that r ." In my view, there is nothing unintelligible, incoherent, or even confusing about saying " A 's reason for Φ ing is that r , but A does not know that r ." If there were anything approaching a paradox here, we

⁶² See Unger [1975: 209]. I expect Williamson [201X], and perhaps also Hyman and Hornsby, would say something similar (perhaps replacing the notion of certainty with that of knowledge).

would expect the conceptual analysis surrounding the cognitive constant to bear that out. But the conceptual analysis does not bear it out.

Perhaps, as with the issue of factivity, there remains an air of paradox about saying that “My reason for Φ ing is that r , but I do not know that r .” I confess that I have a hard time catching wind of this air, but let us grant it anyway. The air of paradox can be explained away, it seems to me, in a manner similar to the air of paradox surrounding claims of the form, “My reason for Φ ing is that r , but not r .” The puzzlement, if there is any, is owed to the fact that, “My reason is that r ” implies that I believe that r . But asserting that “I do not know that r ,” *à la* Moore’s paradox, is at odds with asserting that r . This might be due to the fact that the assertion that r strongly implies that I believe that r , and there is a slight awkwardness (even if it is only slight) about saying that “I believe that r , but I do not know that r .” So, when I say that, “I do not know that r ,” there is some pressure not to assert either that r or that I believe that r ; but if I say that, “My reason is that r ,” this strongly suggests that I believe that r . If there is any puzzlement around remarks of the form, “My reason is that r , but I do not know that r ,” then, it seems to me attributable to a tension found downstream in reports about what I believe to be the case. If this explanation seems to be less satisfying than the explanation of the air of paradox about claims of the form, “My reason is that r , but not r ,” that is likely due to the fact that there is not as strong of an implicature from “I do not know that r ” to “I do not believe that r ” as there is from “It is not the case that r ” to “I do not believe that r .” But to the extent that we feel that the implicature is less strong here, I suspect that we are also less inclined to think there is any air of paradox about claims of the form, “My reason is that r , but I do not know that r .” That being so, even if we think there is *less* of an air of paradox here, my conjecture that the air of paradox is due to the generation of apparently inconsistent belief reports may stand.

All of this leads me to think that the case of RACIST JUDGE (and its variation in which Jerry is not racist) is a clear counterexample to the Knowledge View. It is worth noting that the case also seems to be a counterexample to what I earlier called the Justification View: the view that in order for A to Φ for the reason that r , A must have justification for believing that r . Indeed, one might think that RACIST JUDGE is a counterexample to the Knowledge View in virtue of being a counterexample to the Justification View. Note that this may be so even if, as some believe, that knowledge cannot be analyzed into parts, one of which is justification. Even those who believe that knowledge is primitive or unanalyzable will hold that if A knows that p , then A has justification for believing that p . Since Jerry concludes that HE DID IT! via a biased belief-formation process, his belief that HE DID IT! is not justified. Indeed, it is stipulated in the example that the evidence does not support the claim that Gil is guilty. But again, this seems not to matter to what we can say about what Jerry's reason for acting is. It seems correct to say that Jerry acts for the reason that HE DID IT!, even though he has no good reason for believing that HE DID IT!. Furthermore, even if Jerry *did* have sufficient reason to believe that HE DID IT!, it is not as though this addition would somehow put to rest any suspicions that his reason was not really that HE DID IT!.

I think such cases put a general kind of pressure on the Knowledge View. The pressure is that it is not clear whether any significant, epistemically good connection to a reason will be necessary for acting for that reason. Cases like RACIST JUDGE, though, clearly do not put any pressure on the Belief View. Jerry clearly believes that HE DID IT!, even if he does not know that HE DID IT!. A counterexample to the Belief View would have to be one in which A Φ s for the reason that r but does not believe that r (showing that the Belief View fails to express a true necessary condition on acting for a reason). Here is a case that fits the bill.

UNBELIEVING JUDGE: Judge Jerry is presiding over the sentencing phase of Gil's trial, but has neither delivered the guilty verdict in the previous phase nor come to any belief concerning whether Gil committed the crime in question.⁶³ Two years in prison is the legally appropriate sentence for the crime in question. So, Jerry sentences Gil to two years.

Without doing any sophisticated theorizing, I find it difficult to interpret this story in certain ways. Suppose we add to it, again imagining that a reporter is interviewing Jerry about his perceived harsh sentence. She asks him why he delivered the two-year sentence, and how he knows that Gil committed the crime. We could certainly understand Jerry if he responded, "Look, I am not sure whether HE DID IT!. In fact, I have no view on the matter. Nonetheless, he was convicted of the crime, and I was charged with administering the appropriate sentence for a person convicted of such a crime." However, we *would* have a hard time understanding Jerry if he responded, "Oh, I do not know whether HE DID IT!. In fact, I do not even believe that HE DID IT!. I have no view about that. But nonetheless, I sentenced him to two years for the reason that HE DID IT!." Indeed, even if Jerry *did* say something like this, I think we would be forced, out of charity, to try to interpret him as meaning the first response (which does not imply that he believes that Gil committed the crime). Alternatively, we might try to understand *both* answers as Jerry's telling us that his reason for giving the two-year sentence was *that Gil was convicted of a crime* the legally appropriate sentence for which is two years in prison. In any case, as the audience of Jerry's statements, we would likely go to great lengths to avoid supposing that

⁶³ If for some reason this strains the imagination, we could suppose, for example, that Jerry is taking over for another judge, who already issued guilty verdict for Gil.

Jerry is saying anything of the form, “My reason is that r , but I do not believe that r .” For such an assertion is near incomprehensible, but we feel that we understand perfectly well what Jerry is doing.

Though this is only one case, I think it demonstrates a pre-theoretical bent toward the Belief View. If necessary, we can, as we did before, revisit different conceptual analyses of acting for a reason. I do not think we can intelligibly say, for example, that “Jerry is responding to the consideration that HE DID IT!, but Jerry does not believe that HE DID IT!.” Now, there is nothing peculiar about this case, and it has the ingredients required for concocting a counterexample to the Belief View taken as a necessary condition on acting for a reason. I think this should lead us to believe that it will be at least difficult, if not impossible, to find a counterexample to the Belief View.

The argument of this section is built on two claims. The first is that intuitive conceptual analysis of the notion of *acting for a reason* provides no support for the Knowledge View, but it does lend some support to the Belief View. The second is that a brief examination of cases finds apparent counterexamples to the Knowledge View, but no counterexample to the Belief View — indeed, this examination suggests that a counterexample to the Belief View will not be found. On the basis of these two claims, I think we should conclude that the Belief View is the intuitive view on the issue of the cognitive constraint; and it should be the jumping off point for further theorizing on the issue.

ACTING IN IGNORANCE

With this in mind, the next two sections of this chapter will ask whether there are any arguments compelling enough to draw us away from the Belief View — either toward the Knowledge View, toward some other view discovered along the way, or

toward agnosticism about the cognitive constraint. I will take these suggestions in turn; and in this section, I will address three arguments in favor of the Knowledge View. The first of these amounts to the suggestion that the Knowledge View has a theoretical advantage over the Belief View, since it can explain (while the Belief View cannot explain) why acting for a reason is factive. The second of these comes to the claim that in cases where A does not know that r , we cannot explain A 's Φ ing by supposing that A Φ s for the reason that r . The third, and in my view strongest, argument is that the Belief View does not have theoretical space for what John McDowell has called *acting in the light of a fact*. I think that all three of these arguments fail, and so advocates of the Knowledge View are without compelling argument in their favor. There is no good argument that it is impossible to, as I will say, *act in ignorance*.

What's Truth Got to Do With It?

As I illustrated in the first section, many theorists take the debate over the cognitive constraint to be wrapped up in the debate about whether acting for a reason is factive. This is the debate about, whether, in order for A to Φ for the reason that r , it must be the case that r . The Factive View_M answers this question in the affirmative, while the Nonfactive View_M answers in the negative. Those who endorse the Knowledge View are often tempted to think of the Nonfactive View_M as an opposing view, while those who endorse the Belief View are often tempted to think of the Factive View_M as an opposing view. But, as I demonstrated earlier, the opposition here is not completely straightforward. Neither the truth nor the falsity of the Belief View entails either the truth or falsity of either the Factive View_M or the Nonfactive View_M; and neither the truth nor the falsity of either the Factive View_M or the Nonfactive View_M entails the truth or the falsity of the Belief View. Furthermore, the truth of the Factive View_M does not entail

either the truth or falsity of the Knowledge View; and the falsity of the Knowledge View does not entail either the truth or falsity of either the Factive View_M or the Nonfactive View_M.

As I also illustrated, though, these issues are not entirely unrelated. The truth of the Knowledge View *does* entail the truth of the Factive View_M; and the truth of the Nonfactive View_M entails the falsity of the Knowledge View. Notice, though, that notwithstanding these connections, there are no resources for a deductive argument straight from a position on the factivity of acting for a reason to the truth of the Knowledge View. The argument would have to be more indirect than that. I think there is one such line of reasoning that could be thought to support the Knowledge View. The reasoning goes as follows. Suppose the Factive View_M is true. The Knowledge View contains an explanation of this fact, while the the Belief View does not. So, the Knowledge View has an explanatory advantage over the Belief View, *ceteris paribus*.

This line of reasoning, which I will call the **Argument from the Factivity of Acting for a Reason** (hereafter, “AFAR”), might seem initially plausible or tempting to some. But AFAR suffers from several serious problems which, in my view, prevent it from being an argument we should take seriously. I will address only three. The first, and most obvious, is that the truth of the Factive View_M is contentious; but AFAR only weighs in favor of the Knowledge View if the Factive View_M is true. To the extent that the Factive View_M is controversial, the force of AFAR is to that extent weakened. I argued in the previous chapter that we should think that the Factive View_M is independently implausible. Furthermore, if the Nonfactive View_M is correct, some might even think that the line of reasoning represented by AFAR works *against* the Knowledge View. For, since the Knowledge View entails the Factive View_M, one who finds the Nonfactive View_M

more plausible might have good reason to privilege the Belief View over the Knowledge View.

The second problem with AFAR is that, although the Belief View does not come loaded up, as it were, with an explanation of the truth of the Factive View_M, it is certainly not incompatible with the Factive View_M; nor is it in principle incapable of being supplemented with an explanation of the truth of the Factive View_M, or even with the Factive View_M itself. The possibility of the True Belief View — the mainstream view in action theory — is clear evidence of this. So, even if the Factive View_M were true, and there were some explanatory advantage to be gained by the theory accounting for this fact, it is unclear that this would immediately create an explanatory advantage for the Knowledge View not enjoyed by the Belief View.

The third, and in my view most glaring, problem is that AFAR is predicated on the assumption that the Belief View and the Knowledge View are otherwise equally plausible. But we just saw that the Belief View seems to be the more intuitively appealing view. The fact that the truth of Knowledge View would explain the truth of the Factive View_M is surely not reason enough to abandon the intuitive view — especially since the Belief View is perfectly compatible with both with other theories that might explain the truth of the Factive View_M, and with the Factive View_M itself. It seems, then, that AFAR must be accompanied by other argument for the Knowledge View if it is to have a real impact on the debate.

This is an opportune time to put a bow on a dialectical point I began to develop earlier. In the first section of this chapter, I complained that the issue of the cognitive constraint was problematically intertwined with the issue of the factivity of acting for a reason. We are now in an even better position to see why this intertwining is problematic. The question of the factivity of acting for a reason has only a shaky

connection to the question of the cognitive constraint; and what shaky connection there is does not seem to be one that can be of any help to the Knowledge View. The only clear move from the former issue to the latter is from the truth of the Nonfactive View_M to the falsity of the Knowledge View — but of course, that is of no use to advocates of the Knowledge View. Although the truth of the Knowledge View does entail the truth of the Factive View_M, this connection does not itself provide strong evidence (if any) in favor of the Knowledge View or against the Belief View. Two subsections of this chapter have asked about the issue of the cognitive constraint: What's truth got to do with it? I hope I have at this point illustrated the rhetorical force of this question — or at least that, if there is a serious answer to the question, that answer is: not very much at all.

Knowledge and Reasons-Explanations

So, if there is a good enough argument to persuade us away from the intuitive Belief View, and toward the Knowledge View, the resources for such an argument will not be found exploring the issue of the factivity of acting for a reason. A better argument might be found in the nature of action explanation. We can begin here by recalling the pivotal move in Hornsby's analysis of the well-known case of Edmund, who does not know that THIN ICE: the ice in the middle of the pond is thin. According to Hornsby,

Edmund, then, did not keep to the edge *because* the ice in the middle of the pond was thin... The fact that ice was thin does not explain Edmund's acting, even though Edmund did believe that it was thin, and even though the fact that it was thin actually was a reason for him to stay at the edge.

Hornsby goes on to claim that this directly demonstrates the truth of the Knowledge View. Though the specific structure of this argument is opaque, the path to the conclusion does have two clear steps. One step is from the claim that Edmund does not know that thin ice to the claim that Edmund did not keep to the edge *because* thin ice.

Another step is from this claim to the claim that Edmund did not act for the reason that THIN ICE. This is supposed to lead us to the Knowledge View, since Edmund's not knowing that THIN ICE appears to prevent him from acting for the reason that THIN ICE. The implication is that any reasons-explanation of an action entails that the agent performing the explained action has knowledge of the reason explaining the action.

Let's take a closer look at this kind of reasoning. The second step here might seem sensible in view of our now-familiar axiom of

REASONS-EXPLANATIONS: a motivating reason is the kind of thing that can play a role in explaining an agent's action (when the action is one the agent does for a reason).

This axiom reveals some wisdom in supposing that there is a connection between the claim that $A \Phi$ d for the reason that r and the claim that $A \Phi$ d because r . A telling commonality is that both claims can serve as good answers to the question, "Why did $A \Phi$?" Why does this matter? A natural thought, we have seen, is that answers to "Why?" questions are explanations. Another natural thought, then, might be that both of these claims are of the right kind to be explanations of A 's Φ ing. In wondering why people do the things that they do, we often prefer that the answers to the relevant "Why?" questions offer the reasons for which the agents did what they did. The claim that $A \Phi$ d for the reason that r and the claim that $A \Phi$ d because r both seem capable of accomplishing this. Indeed, we might even take these two claims to be, at bottom, the very same explanation articulated in different ways. If true, this would establish quite an intimate connection — one even more intimate than sharing the common feature of simply offering *an* explanation of A 's Φ ing. I think it not outlandish, then, to make the second step of

Hornsby's argument for the Knowledge View here. It is not obviously misguided to think that if Edmund could not have acted *because* THIN ICE, then he could not have acted *for the reason that* THIN ICE.

It is not *obviously* misguided, but it is misguided — or so I argued in the previous chapter, and will continue to argue in the next chapter. We simply cannot read the facts about the availability of claims of the form, “ $A \Phi d$ for the reason that r ” off of the availability of claims of the form, “ $A \Phi d$ because r .” So any argument in the spirit of Hornsby's, I think, needs to face up to the points of the previous chapter. At this juncture, though, I want to focus on the first step of Hornsby's argument. Why should we expect there to be a connection between claims about A 's cognitive relation to r and explanatory claims concerning A 's Φ ing? Consider the following case.

INCONSIDERATE EDMUND: Edmund is skating on a pond, keeping to the edge as he skates. You are on a nearby hill, observing the skaters on the pond. You notice that Edmund is skating only on the edge of the pond. As you prepare to skate on the pond, you approach Edmund and ask him whether the ice in the middle of the pond is thin. Edmund responds by saying that he had not considered whether the ice in the middle of the pond was thin.

Now, these details do not explicitly disclose Edmund's reason for skating near the edge of the pond. But it would be quite unnatural, I think, to describe Inconsiderate Edmund as a case in which Edmund skates near the edge *for the reason that* THIN ICE. Suppose you continue the conversation with Edmund: “Oh... well, then, why are you skating only near the edge of the pond?” — to which he replies that THIN ICE. You would hardly receive this answer. Indeed, I think any of us would struggle to even understand this

answer, and that is because such an answer is at odds with Edmund's previous claim that he had not considered whether THIN ICE. Here we can recall some of our earlier analyses of acting for a reason. It strains the imagination to suppose that Edmund could have responded to a consideration that he did not consider; that he treated as a premise in deliberation a claim that he did not consider; and so on. It is difficult to suppose that that one could be motivated by a consideration to which one bears *no* cognitive relation at all.

So, I think we should allow that, in general, the *kind* of move made in the first step of Hornsby's argument for the Knowledge View is legitimate. The explanatory role of motivating reasons clearly limits what we can say about agents' cognitive relations to their reasons, and vice versa. The question, then, is whether the *particular* move that Hornsby makes is legitimate; the question is how, specifically, the explanatory role of reasons and the cognitive constraint bear on one another. Hornsby's argument gives us a well-defined target. Is the connection here such that if *A* does not know that *r*, then we can conclude that *A* could not have Φ d because *r*?

I think not; let me explain why. The argument proceeds immediately from the idea that Edmund does not know that THIN ICE to the claim that Edmund does not keep to the edge *because* THIN ICE. But it is not at all clear what the connection here is supposed to be. It is certainly not apparent that the claim that Edmund does not know that THIN ICE *entails* the claim that Edmund does not act because THIN ICE. Indeed, there seem to be obvious counterexamples to the view that claims of the form "*A* does not know that *r*" entail claims of the form "*A* does not Φ because *r*." Suppose, for example, that Edmund only skates near the edge because he has an extreme fear of drowning; he would skate all over the pond if not for this fear. Suppose also, though, that Edmund has repressed any memory of the traumatic near-death experience which brought about this fear, such that he does not even know that he has an extreme fear of drowning. He is not the slightest bit

aware of this fear. This seems to make no difference to whether we can say, truly, that Edmund skates near the edge because he has a fear of drowning. The very same point applies to the claim that “What explains Edmund’s skating near the edge is that he has an extreme fear of drowning.” The fact that Edmund does not know that he has this fear does not preclude the truth of this explanatory claim.

Now, one might reasonably object here that the explanatory claim targeted by Hornsby’s argument ($A \Phi$ because r) does not employ just any sense of ‘because’ or ‘explains’. Rather, it employs the *rational* sense. This is correct, but it is not clear how this kind of argument for the Knowledge View can make use of this idea. Suppose the argument goes as follows:

- [1] If A does not know that r , then A cannot Φ because r .
- [2] If A cannot Φ because r , then A cannot Φ for the reason that r .
- [3] Therefore, if A does not know that r , then A cannot Φ for the reason that r .

We have already seen, in chapter one (and we will see even more so in chapter three), how claim [2] here might be challenged. We are presently focused on claim [1]. The criticism is that it is not obvious why anyone should believe [1]. It is hardly self-evident, and there does not seem to be any entailment relation that could underpin it. Indeed, it seems to suffer counterexamples of the sort just mentioned. So, we have come to a strategy for fending off such counterexamples. The strategy is to point out that the counterexamples trade on one sense of ‘because’, but the argument for the Knowledge View employs another sense — the *rational* sense. But this clarification is of no use. For what someone like Hornsby might mean by “rational sense” here is that “because” could

be glossed as “for the reason that.”⁶⁴ There is nothing wrong with this claim in itself; indeed, I should like to allow that this is one legitimate use of “because.” But allowing this glossing to qualify claim [1] results in the argument’s being clearly question begging: claim [1] would then be identical to claim [3], and claim [3] is exactly what is at issue.

So, while it seems sure that there is a general connection between the cognitive constraint and reasons-explanations of actions, the particular connection that this Hornsby-style argument seems to rely on is dubious. Without any further argument or explanation as to why this connection holds, we should reject this argument for the Knowledge View. In particular, we should require of advocates of the Knowledge View an argument for the view that the claim that A does not know that r secure the claim that A does not Φ because r (or that A ’s Φ ing cannot be explained by r , or some other such).

Acting in the Light of (What is Believed to Be) a Fact

Of those who have tried to provide an account of this connection, I think John McDowell has made the most compelling case. In trying to explain how he thinks the Hornsby-style argument is supposed to work, McDowell says that, in cases like that of Edmund,

... we know that [the agent] takes the relevant fact to obtain, and we know that it does obtain. But in knowing only that much, we do not know whether its obtaining is anything but a happy accident in relation to her cognitive position. [This] version of acting in the light of a fact is not sensitive to the difference between cases in which the obtaining of the fact is, and cases in which it is not, a mere happy accident in relation to the agent’s cognitive position... If, but only if, the obtaining of the fact by virtue of whose obtaining the relevant belief is true is *not* a happy accident in relation to the agent’s cognitive position, we can say that the fact itself is exerting a rational influence on the agent’s will; we can say that in doing what [he] is doing the agent is responding rationally to the fact itself.⁶⁵

⁶⁴ Indeed, Hornsby [2008: 250] says as much.

⁶⁵ See McDowell [2013: 16-17].

For my part, this is the most promising line of argument in favor of the Knowledge View in the literature. It makes it clear why one might think that if A does not know that r , A cannot Φ for the reason that r . The idea here seems to be that if A does not know that r , then r cannot be used in a reasons-explanation of A 's Φ ing; and the idea *there* is that that if A does not know that r , then A 's cognitive relation to r might be a mere accident. But if A 's responding to r is a mere accident, McDowell seems to suggest, we cannot explain A 's Φ ing by saying that A responded rationally to r (or that A Φ d for the reason that r).

Though what McDowell says here is an improvement on the original Hornsby-style argument, I think these remarks do not do enough to clarify or make successful the argument for the Knowledge View as it has so far been discussed. I do think, though, that McDowell's remarks allow us to make room for insights captured by the Knowledge View in a way that does not impugn the Belief View. I will begin by focusing on what I take to be two fatal flaws in the argument if it is taken to support the Hornsby-style Knowledge View. The first is that it is simply not clear what is meant when McDowell says that if the agent does not know that r , then the agent cannot be responding to the "fact itself." For this notion to pull any real weight in the argument, there must be some deep difference between cases in which A Φ s in response to the consideration that r and it is a fact that r , and cases in which A Φ s in response to the-fact-that- r . For if there is no deep difference here, the Belief View could dodge the argument simply by endorsing the Factive View_M; so the argument would not really be an argument in favor of the Knowledge View *per se*.

It is worth keeping in mind Hornsby's own remark about how this sort of argument is supposed to work.

Edmund is a familiar sort of character in epistemology. Such characters are usually used to show that justified true belief is not sufficient for knowledge. Edmund here is used to show that someone's having a true belief (even a justified one) which explains their acting is not sufficient for them to have acted for an (F)-type reason.⁶⁶

It is clear, then, that this kind of argument for the Knowledge View is not one which is supposed to be capable of being sidestepped by endorsing the Factive View_M (or the Justification View). Indeed, per McDowell,

It does not add to the explanatory power of explanations given... if things are as the agent relevantly takes them to be. If things are that way, that is just an extra fact about the situation; the action would have had the same intelligibility... if the agent had been wrong in taking things to be that way.⁶⁷

The deep gulf between responding to the consideration that *r* and responding to the-fact-that-*r*, then, cannot be bridged by adding to the former that *r* is a fact.

But what, exactly, could the deep difference here be? The difference cannot be located in the character or content of the consideration to which the agent responds, since it is the same consideration — THIN ICE, for example — in both kinds of cases. Neither can the difference be in the agent's response to the consideration: for in both cases, the agent might respond in the same way (by skating near the edge of the pond, for example). But the consideration and the response to the consideration seem to be the only two *relata* involved in Φ ing for the reason that *r*. If both *relata* are the same in both kinds of cases, it is not clear what the important difference here could be; there certainly does not seem to be any difference which could make one of these a case of Φ ing for the reason that *r* while precluding the other from being a case of Φ ing for the reason that *r*.

⁶⁶ See Hornsby [2008: 251].

⁶⁷ See McDowell [2013: 19].

One could at this juncture add a new detail: in the second case where A Φ s in response to the-fact-itself, A Φ s in response to the *known* fact that r — and this is what distinguishes cases of the second kind from cases of the first kind. Another way of putting this is that, while perhaps the two-place relation of Φ ing for the reason that r does not imply knowledge that r , the three-place relation of A 's Φ ing for the reason that r *does* imply such knowledge (on the part of A). But such a move would be infelicitous. For what we were looking for in the first place was a distinguishing feature of *responding to the-fact-itself* such that the agent should have to know the fact in order to achieve this feat. It clearly will not do, then, to stipulate that in cases of *responding to the-fact-itself*, the agent knows the fact, and that this is the distinguishing mark of the phenomenon. The inappropriateness of such a move is even more apparent when we consider the broader context. We are trying, *à la* Hornsby, to find a significant difference between cases in which A has, say, a justified true belief that r , and those in which A knows that r , such that the latter could count as acting for a reason while the former could not. Such a difference might give us reason to prefer the Knowledge View as a cognitive constraint on acting for a reason. But if the suggestion is that cases of acting for a reason are distinguished as cases of responding to the-fact-itself, then the phenomena of responding to the-fact-itself had better not itself be underpinned only by or grounded only in the fact that the agent knows the fact in question. For then the overall argument would threaten to collapse into the suggestion that only in cases where A knows that r does A act for the reason that r , because only in those cases does A know that r .

So, although McDowell seems to make some importance out of the notion of responding to the-fact-itself, it is not clear what could distinguish this notion in a way advantageous to the Knowledge View. The second flaw in McDowell's argument is not an unclarity but rather an intuitively false premise. The premise is that in order for A to

respond rationally to the fact that r , A must know that r . This seems to me clearly false upon the briefest examination of examples. For consider the case in which Edmund has good reason to believe the fact that THIN ICE, and then responds to this fact by keeping to the edge of the pond when he skates. Now, if one takes it that the ice in the middle of the pond is thin, it is surely rational to keep away from the middle of the pond. In skating near the edge of the pond, then, Edmund is responding to the consideration that THIN ICE, his response is rational, and it is a fact that THIN ICE. So it seems undeniable to me that Edmund is responding rationally to the fact that THIN ICE. Furthermore, we can say this without saying anything about whether Edmund *knows* that THIN ICE. Indeed, we could even add to the story that he does not know that THIN ICE. This makes no difference; we can still say that Edmund responds rationally to the fact that THIN ICE. But if this is right, then, one of the crucial points in McDowell's argument is false.

At this point advocates of the Knowledge View might object that I have misunderstood their position, and the arguments motivating it. Here is Maria Alvarez, aptly summarizing things:

If there is a fact in virtue of which her belief is true, then she acts in the light of that fact, or is guided by that fact, only if she knows that fact. If the agent does not know the fact, then we cannot say that she was guided by it (Hyman), or that she was responding to it rationally (McDowell). If the agent does not know the fact, the argument goes, the relationship between the agent's acting as she did and the fact is fortuitous, a matter of luck or coincidence, and hence not sufficient for the fact to be her reason for acting. And this, they argue, is so even in cases where an agent acts motivated by a belief that is both true and justified...

Their point is that there is a notion of acting for a reason—arguably, the central notion—that involves the idea of being guided by a fact. This notion requires not mere belief but knowledge of the fact that is a reason.⁶⁸

⁶⁸ See Alvarez [2016: 28].

There are two suggestions here. The first is that any accidental connection between the motivating consideration and the action impugns that consideration's status as a reason for acting. The second is that there is a distinct and important notion — that of acting in the light of a fact — which is not captured by the Belief View.

The second point can be dealt with using resources already available to us. We have seen that there is nothing in the conceptual analysis of Φ ing for the reason that r which suggests that r must be a fact. Now, once we are rid of this assumption of factivity, what reason do we have to suppose that there is a distinct and central variety of acting for a reason, *being guided by a fact* — and what is its distinguishing feature? The Belief View allows that, often when A Φ s for the reason that r , it is the case that r . In such a case, A 's Φ ing is motivated by the (true) consideration that r . What is the difference between this and A 's Φ ing in the light of the fact (or A 's being guided by the fact) that r ? One option is to say that there is something special about responding to r , the-fact-itself (to be distinguished from responding to r , which is a fact). But we have just seen that it is opaque what the distinguishing feature of that phenomenon could be. Another option is connected to the first point Alvarez observes. The option is to say that there is a variety of Φ ing for the reason that r such that, when it is the case that r , the connection between r and A 's Φ ing is not “fortuitous” (or, as McDowell would say, a “happy accident”).

One question that immediately arises concerning this second option is: in what sense is the connection not fortuitous, such that it has remedied a fortuitousness in the case where A does not know that r ? We can approach the question by again noticing what can be said using only the resources available to the Belief View. On this view, Edmund may skate near the edge of the pond for the reason that the ice in the middle of the pond

is thin, so long as he believes that the ice is thin. Now, once we grant that Edmund takes it that THIN ICE, there is one clear sense in which the connection between *the consideration* that THIN ICE and Edmund's skating near the edge of the pond is not a mere happy accident. The sense is this: given that THIN ICE is a thing considered by Edmund (and taken to favor skating near the edge of the pond), it is no accident that he skates as he does. Evidence for this is that we could satisfactorily explain his skating near the edge by saying that "In skating that way, he is responding to the consideration that THIN ICE."

Advocates of the Knowledge View may respond, "Ah — but it remains a happy accident, *given only that* THIN ICE, that Edmund skates as he does. Evidence for this is that we could not explain his skating by saying that he is skating that way because THIN ICE." I make two points in response. The first is that, if we allow that "because" may here be glossed as "for the reason that THIN ICE," then this line of argument begs the question in favor of the Knowledge View in the manner observed in the previous section. On the other hand, if we say that "because" here is the more general "because," then (as we also saw in the previous section) the idea that we cannot say that Edmund is skating as he is *because* THIN ICE seems to endorse a general view about the connection between knowledge claims and because claims that is subject to counterexample.

The second point of response here is that, even if adding that Edmund knows that THIN ICE in some sense makes the connection between THIN ICE and his skating near the edge less fortuitous, it is still unclear in what sense this is so. Even if Edmund knows that THIN ICE, it may still remain a happy accident, *given only that* THIN ICE, that Edmund skates as he does. Perhaps, for example, he only knows about the ice because he decided to go out in the first place, but he did *that* only on a whim. Perhaps he only inspected the ice because Edna, who is known to play tricks on him, said that the ice was thin. Perhaps Edmund thinks that THIN ICE favors his going to the center of the pond and stomping

vigorously. There is an unending set of such circumstances that could seemingly reintroduce any possible fortuitousness supposedly dissolved by the introduction of knowledge on the part of the agent. So on this front we should like to hear much more from one offering the Hornsby-McDowell-style argument in favor of the Knowledge View.

At this point the dialectic has perhaps been muddled; allow me to try to lend some clarity. We began this section asking whether there is compelling argument to reject the intuitive Belief View in favor of the Knowledge View. We noticed that the Hornsby-style argument seemed to assume a connection — an entailment, perhaps — between the claim that A does not know that r and the claim that A does not Φ because r . We saw that, insofar as this argument relies on a claim about a connection between knowledge and reasons-explanations of action, it either begs the question in favor of the Knowledge View or contains an implausible premise. We then encountered a suggested improvement on the argument from McDowell, which seemed to take us away from the issue of explanation and back toward the conceptual analysis of acting for a reason. We saw that this argument seems to rely on a dubious distinction between responding to the consideration that r (which is a fact) and responding to r , the-fact-itself. Finally we arrived at an attempt to legitimize this distinction via the notion that when an agent responds to the-fact-itself, the connection between the fact that r and A 's Φ ing is not a mere happy accident. But we were immediately confronted with the question of what is the sense in which the connection is not a happy accident; and we struggled to answer that question in a way that gave advantage to the Knowledge View.

The overall point, then, has been that advocates of the Knowledge View seem not to have sufficient resources for clearly demarcating the supposedly central notion of *acting in the light of a fact* without begging the question in their own favor. Unless these problems

can be worked out, I think we should say that even McDowell's improvements are not enough to create a successful argument for the Knowledge View. I am doubtful that any forthcoming amendment to either AFAR, the Hornsby-style argument depending on a connection between knowledge and action explanation, or the McDowell-style argument depending on carving out a distinct notion of acting in the light of something known will succeed. This is because, at bottom, each of these arguments is committed to implausible claims about either the nature of action explanation or the concept of acting for a reason.

ACTING IN UNBELIEF

Aside from a bit of conceptual analysis and casuistry, I have not yet provided any extended defense of the Belief View. I have only argued that the Belief View is the intuitive view, and that the arguments usually given for the Knowledge View should not persuade us that it is preferable to the Belief View. In arguing that the Belief View is the intuitive view, I attended to the case of UNBELIEVING JUDGE. I argued that this case suggests that whatever the correct cognitive constraint is, it cannot be weaker than belief. I take it that this is a point of unanimous agreement in the literature. No one, to my knowledge, has argued for anything other than the Belief View or the Knowledge View, and no one has so much as suggested that the constraint might be any weaker than belief.⁶⁹ In this section, though, I want to address an apparent counterexample to the Belief View. I will argue that although the example is cause for some initial concern, it can be shown not to be a counterexample to the Belief View. Furthermore, I will argue that there is very good theoretical reason to resist any apparent counterexample to the Belief View, as any theory of acting for reasons which weakens the cognitive constraint beyond belief will have serious shortcomings.

⁶⁹ Jonathan Dancy has told me that he has occasionally wondered about this, though not in print. David Sosa, too, has registered his doubts about the Belief View in conversation.

The case I am interested in is that of

CURBSIDE CHRIS: Chris is waiting to be picked up from his hotel by his friend Larry. Chris and Larry were once the best of friends, but unfortunately, Chris's trust in Larry has gradually eroded; Larry has become unreliable to the point that Chris no longer believes that Larry will do as he says. But Larry, in a last-ditch effort to restore Chris's trust in him, has promised to provide transportation for Chris on this particular occasion. Chris has packed his bags and is waiting curbside to be picked up by Larry, but he does not actually believe (or disbelieve) that Larry is coming to pick him up. When the hotelier asks Chris why he is waiting curbside, he responds by saying that his friend Larry is coming to pick him up.

What should we say about Chris?⁷⁰ If any of us were the hotelier, we would certainly think that Chris was waiting curbside *for the reason that* his friend Larry is coming to pick him up. Even if Chris disclosed to us that he did not really believe that Larry was coming, we would still be able to make sense of his response. If he told you the whole story of CURBSIDE CHRIS, for example, and one of the other employees at the hotel asked you what was going on with Chris, I think you would not struggle to come up with something like, "He is waiting for his friend Larry to come pick him up." This would surely be an informative answer, and one that we should count as a good reasons-explanation of what Chris is doing.

⁷⁰ In formulating this kind of case I am indebted to Chris Simpson, Brian Cutter, and Dan Bonevac. Bonevac put to me a case involving a baseball player who has been called out trying to run to the next base. Though the team manager does not believe he was out, he calls the next batter, plausibly for the reason that the other player is out.

I find the case of CURBSIDE CHRIS to be a fairly compelling counterexample to the Belief View. It might appear sensible — even correct — to say that Chris is waiting for the reason that Larry is coming to pick him up, even while holding that Chris does not believe that Larry is coming to pick him up. If this is correct, then the cognitive constraint on acting for a reason is weaker than belief. Perhaps, for example, the correct constraint is one of lacking disbelief: in order for A to Φ for the reason that r , A must at least not believe that *not* r . Perhaps the correct constraint is a matter of credence, confidence, or some estimation of likelihood with respect to the motivating consideration. Or perhaps the correct constraint involves a disjunction of more complex attitudes that are distinguishable from belief. Perhaps, for example, the case of Curbside Chris only makes sense because, while Chris does not believe that Larry is coming to pick him up, he *hopes* that he is; and he is, as it were, acting on that hope.

While I think all of this has some plausibility to it, I think we can give a satisfactory account of the case of Curbside Chris without abandoning the Belief View. Two accounts that naturally suggest themselves are that Chris is waiting curbside for the reason that Larry *might* pick him up, and that *Larry said* he would pick him up. Both of these claims are consistent with Chris not believing that Larry will pick him up, and they also make for perfectly intelligible explanations of what Chris is doing. I think that in general, though, we should hesitate to make such moves. We are not commonly incredulous with respect to the reports others give us about what their reasons for acting are. Furthermore, one could sensibly ask why one should allow such a move here, but not for the cases of RACIST JUDGE, UNBELIEVING JUDGE, INCONSIDERATE EDMUND, and so on. So I think I should say something to motivate this treatment of the case.

Consider familiar cases in which we follow directions from strangers. In these familiar cases, we sometimes form beliefs about how to get to our destinations on the basis

of testimony. But often times we form no such belief; we only get so far as the belief that, for example, *she said* that it was just around the corner. To make a clear connection to the case of CURBSIDE CHRIS, consider the following case.

TRICKY TOWN: Travis is in a new town, and he is looking for the local grocery store. He has had no luck finding it, so he resolves to ask a stranger how to get to it. Travis has heard, however, that the current residents of this town resent visitors and new residents; and they have colluded to make things unpleasant for such people in a number of ways, chief among them being to give such people bad advice and misleading directions. Nonetheless, Travis has resolved to ask a stranger for directions to the local grocery store. She tells him that it is just around the corner. Travis then goes around the corner.

Now, sometimes, we trust the testimony of strangers to the point that we come to believe what they say. But TRICKY TOWN seems obviously to be a case in which one would not do this. Travis would likely *not* form the belief that the store is just around the corner; he would more likely restrict himself to forming the belief that *she said* that the store is just around the corner. Now, for what reason does Travis go around the corner? It seems to me that, even if Travis says that “The grocery store is just around this corner,” we should not take this to be a precise specification of his reason. This is because Travis is plainly *not* acting in the light of the consideration that the store is just around the corner — this consideration is not what motivates Chris to go around the corner. Rather, it seems that the consideration he is responding to is that *she said* that the store is just around the corner, or perhaps that the store *might be* around the corner. We need not settle on the positive account as to what Travis’s motivating reason is. All that is required for my

present point is that there is nothing forcing us to say that Travis's reason is that the store is just around the corner, and also that it is plausible that his reason could just as well be that she said that the store is around the corner. These possibilities are made salient by the fact that Travis explicitly doubts the truth of the claim that the store is around the corner.

It seems to me that Travis and Chris are in similar situations, since Chris is also acting on the testimony of someone he does not take to be reliable. The only difference is in what is the most natural way for these two agents to specify their reasons in conversation. So I think we can plausibly say that Chris, like Travis, does not act for the reason that r , but rather his reason is something like *it might* be the case that r , or *he said* that r .

I hope that this is enough to illustrate how proponents of the Belief View can plausibly deal with a case like CURBSIDE CHRIS. To shore up this position, I want to point out that there are cases in which not believing that *not* r , for example, seems quite far off the mark from enabling the agent to Φ for the reason that r . The earlier case of INCONSIDERATE EDMUND demonstrates this. It is nearly unintelligible for Edmund to say, for example, "The consideration that the ice in the middle of the pond is thin never occurred to me, and I am keeping to the edge for the reason that the ice in the middle of the pond is thin." Clearly, in order for Edmund to act for the reason that THIN ICE, his cognitive relation to THIN ICE must be stronger than this. How much stronger? It seems to me that belief is a good candidate; it is at least not clear to me what weaker relation might get the job done.

Where does all of this leave us? I hope to have established several points. First, the Belief View is the intuitive view of the cognitive constraint on acting for a reason. Second, the best arguments in favor of the Knowledge View are not strong enough to persuade us

that knowing that r (or any other cognitive relation stronger than belief) is necessary for Φ ing for the reason that r . Third, the Belief View has the resources to address examples that might initially seem to show that belief that r is not necessary for Φ ing for the reason that r . Fourth, the constraints suggested by apparent counterexamples seem not to be sufficient cognitive conditions for Φ ing for the reason that r . Finally, outside of these (merely) apparent counterexamples, believing that r does seem a good candidate for what is necessary and sufficient (*qua* cognitive constraint) for Φ ing for the reason that r . I think that these considerations should lead us to believe that, while it is possible to (as I say) act in ignorance, it is not possible to act in unbelief.

ARE REASONS BELIEFS?

I now want to, supposing that the case has been made for the Belief View, consider what implications this might have for other theories of reasons. To keep things brief, I want to consider just one quite popular position in the theory of reasons which seems to me to be at odds with the Belief View. Now, it may seem obvious, after the discussion of this chapter, that agents who act (or believe) for reasons must bear some cognitive connection to those reasons. Though we have seen that some think that the specific cognitive connection required is knowledge, most everyone holds that at least belief is required. A superficially similar view has been put forth — and gained significant favor — about the ontology of reasons. This view, often called Psychologism, holds that the reasons for which agents act are psychological states of (or psychological facts about) those agents. I want to conclude this chapter on the cognitive constraint by showing that there is a deep tension between the Belief View and Psychologism, such that one cannot plausibly subscribe to both. I proceed to argue that the best way to resolve this tension is

to reject Psychologism. In so doing, I hope to be offering a novel argument against this well received theory.

The course of the section is as follows. First, I recount the basic picture and appeal of both the Belief View and Psychologism. Second, I illustrate the tension between these two views by demonstrating that any theory of reasons that incorporates both of them will be committed to an implausible analysis of acting for a reason. In the third section, I make the case for resolving the tension by rejecting Psychologism. What ultimately emerges is that Psychologism is at odds with the thought that there is *any* cognitive constraint on acting for a reason. Rather than give up that extremely plausible thought, I conclude in the fourth section, we should reject Psychologism.

Two Common and Appealing Views

As I said, it seems obvious that agents who act for reasons must bear some significant cognitive connection to those reasons; there must be some cognitive constraint on acting for a reason. Put differently, we can say that there must be a significant answer to the question: what cognitive relation must some agent A stand in to some reason r in order for A to perform some action Φ for the reason that r ? We have seen that the two main competing answers to this question are as follows.

Belief View: In order for A to Φ for the reason that r , A must believe that r .

Knowledge View: In order for A to Φ for the reason that r , A must know that r .

There has been much recent discussion between proponents of these two views of the cognitive constraint, and I have spent the bulk of this chapter arguing that the Belief View is the more plausible cognitive constraint on acting for a reason. But for present purposes

we can leave this debate undecided; my argument in this section will succeed whether it is the Belief View *or* the Knowledge View that is ultimately correct. Since it is the Belief View that is much more widely received, though, I will center my discussion around it.

The idea that the cognitive constraint on acting for a reason must be at least as strong as belief has the status of something like an assumption among theorists.⁷¹ That being the case, not much has been said explicitly in its favor (except in addressing the competing Knowledge View). We were able to see some of the intuitive appeal of the view earlier, though. A common analysis of acting for a reason is *acting in the light of a consideration*. Put less metaphorically, we can say that acting for a reason is a way of *responding to a consideration*. For those who do not find it immediately transparent that one cannot respond to anything that one does not believe to be the case, we saw, we can attend to cases like that of

UNBELIEVING EDMUND: Imagine that you are about to go skating, and you see Edmund skating on the nearby pond. You notice that when he skates, Edmund keeps to the edge of the pond. As you are about to begin skating, you ask Edmund why he is skating only around the edge of the pond. He responds by saying, “My reason is that the ice in the middle of the pond is thin.” When you thank him for warning you about the thin ice, he says, “Oh — I was not saying that the ice in

⁷¹ Although this is widely assumed, there might be ways in which it is contentious. For example, Nomy Arpaly [2003: 37-65] may be committed to denying the claim — but only by taking a somewhat idiosyncratic stance on what we might call “Freudian” cases: cases in which the agent’s motivating reason for acting is obscure to the agent. Notice, though, that the conceptual analyses I provide here put such cases to the side (since *A* cannot respond to *r*, in the relevant sense, without taking it to be the case that *r*). Furthermore, the mere existence of Freudian cases does not challenge the Belief View; that *A* does not know or believe that *A*’s reason is that *r* does not show that it is not the case that *A* believes that *r* (or that *r* is *A*’s reason for acting).

the middle of the pond is thin. In fact, I do not even believe that it is. Nonetheless, that is the consideration I am responding to in skating near the edge of the pond.”

There seems to be some tension in the reports Edmund provides — one as to the consideration he responds to in acting (and to his reason for acting), and the other as to what he believes to be the case. Our puzzlement here is plausibly attributed to a tacit assumption that there is a tight conceptual connection between responding to the consideration that r and believing that r . Similarly, it is almost Moore-paradoxical to say, “My reason for acting is that r , but I do not believe that r .” If the Belief View is true, then each of Edmund’s claims cannot be true at once; and so something like a tacit endorsement of the Belief View easily explains our puzzlement at Edmund’s remarks.

So the Belief View, even if not ultimately correct, at least has some strong initial appeal. My initial aim here is to show that this view is at odds with another view, which has been called

Psychologism: whenever $A \Phi$ s for the reason that r , r is a psychological fact about A .

The main competitor to Psychologism, **Non-Psychologism**, is just the denial of Psychologism. Psychologism is a theory of reasons with a large and enthusiastic following, in no small part due to the work of Donald Davidson.⁷² It is worth noting, though that if

⁷² See Davidson [1980]. Too many to name subscribe to the view, but Smith [1998; 2010], Alfred Mele [1997; 2013], and Michael Bratman [1999] are perhaps also worth mentioning. John Turri [2009] has explicitly defended Psychologism about reasons for belief, while Clayton Littlejohn [2012] seems to resist it.

I count Alvarez [2010: 131-135], Dancy [2000: 98-120], Hornsby [2008], Hyman [1999: 442-446; 2011], and Raz [1975] among those Non-Psychologists who have given compelling arguments against the view, but I will not revisit those arguments here. In epistemology, Enoch [2010: 981-986] and Williamson

Psychologism is true, much of our ordinary talk around reasons is, strictly speaking, incorrect. To see why, consider the following case.

FLORENCE'S FLOWERS: Florence's mother is in a funk. To cheer her up, Florence decides to buy her some flowers. So, she goes to the store to get some flowers. Taking it that roses are her mother's favorite flowers, she picks out some roses. As she checks out, the cashier asks her, "Why are you buying roses — what's the occasion?" Florence responds by saying, "I am trying to cheer up my mother, and roses are her favorite flowers."

I think we all find it natural to interpret this story as one in which Florence buys roses for the reason that *roses are my mother's favorite flower*, or perhaps, that *roses will cheer up my mother*. But Psychologism would say of Florence that she cannot be buying flowers for either of these reasons. For such claims incorrectly imply that Florence's reason is a state of affairs in the (external) world. Rather, Psychologism tells us that Florence is acting for the reason that, for example, *she believes* that roses will cheer up her mother. I say "for example" here because, strictly speaking, all that is necessary for Psychologism to say in this example is that Florence's reason is *some* psychological fact (or state) or another. It need not involve belief. It could be, for example, that Florence *wants* to cheer up her mother. I center my discussion around belief because it is (rightly) thought to be the most plausible candidate. It may help to note that the more traditional Psychologism, according to which reasons are belief-desire pairs, will include the "B-Psychologism" I focus on shortly. In any case,

[2000; 201X] have also expressed Non-Psychologistic sympathies. know or believe that *A*'s reason is that *r* does not show that it is not the case that *A* believes that *r* (or that *r* is *A*'s reason for acting).

according to Psychologism, Florence's reason is not a state of the world, but rather a state of her psychology.

Offense to our natural ways of speaking notwithstanding, Psychologism itself (like the Belief View) also possesses some intuitive appeal. For we could just as easily imagine FLORENCE'S FLOWERS as a story in which Florence gives her reason by saying, "Well, I am trying to cheer up my mother, and I think that roses will make her happy." Though it is natural to say that Florence's reason is that roses will cheer up her mother, it is perhaps just as natural to say that Florence's reason is *that she believes* that roses will cheer up her mother.

Both the Belief View and Psychologism enjoy widespread support among theorists of reasons. The advantages and disadvantages of both views have been well documented and discussed by many, and I shall not attempt at this point to thoroughly present anything like a complete case for or against either view. My purpose here is to develop a novel, and in my view devastating, argument against Psychologism. The argument is fairly simple: I will argue that if the Belief View is true, then Psychologism is false. Before moving to the substance of the argument, I want to raise a dialectical query. Given that these two theories enjoy the widespread support that they do, one might think that my strategy should be met with some suspicion. It might be thought puzzling, if these two views are so widely discussed and received, that my argument against Psychologism has not yet been raised.⁷³ But I think that there is a good explanation for this state of affairs – and one which will help to elucidate my central argument. The explanation is that it might initially seem that Psychologism either includes, is a variant of, or otherwise agrees with the Belief View. As Alvarez puts it,

⁷³ Although the thought that something like the Belief View and Psychologism are opposed has perhaps been in the air — see, for example, Alvarez [2010: 124-130] and Enoch [2010: 981-986].

Most if not all accounts of acting for a motivating reason require as a condition that the agent be in some kind of epistemic relation to the reason that motivates her. And we also saw that a widespread view is that this epistemic relation is one of belief: for an agent to act for the reason that p , the agent must believe that p . It is this thought that led many to endorse the view that reasons are mental states (often as part of the “desire-belief” conception of reasons for action described above).⁷⁴

Now, where Non-Psychologism says that $A \Phi$ s for the reason r , Psychologism will say that $A \Phi$ s for the reason that A believes that r . Consider, then, the superficial similarity between the Belief View and Psychologism when they are expressed slightly differently from before:

B-Psychologism: Whenever $A \Phi$ s for the reason that r , r is the fact that A believes that p .⁷⁵

Belief View: Whenever $A \Phi$ s for the reason that r , r is something A believes (namely, p).

It is not difficult to see how these two views could appear to be in agreement. In FLORENCE’S FLOWERS, one might think, both B-Psychologism and the Belief View centrally feature Florence’s belief that ROSES: Roses will cheer up my mother. They will both say, for example, that Florence’s belief that ROSES is essential to her acting as she does. They both also imply that any good reasons-explanation of Florence’s buying the roses relies on the supposition that Florence believes that ROSES. Without too much

⁷⁴ See Alvarez [2016: 27].

⁷⁵ There is, of course, the alternative rendering: “ r is A ’s belief/believing that p .” This option might be preferred by those Psychologists who insist that it is the agent’s psychological *state* — not a *fact about* what the agent’s psychological state is — that is the reason for which the agent acts. I think, however, the problems that I will introduce for Psychologism are only augmented by this statist articulation of the view. This is because, at bottom, the problem is that Psychologism must somehow accommodate the thought that motivating reasons are things capable of being believed to be the case; but it is much more awkward to see how a psychological state *itself* (rather than the content of that state, or the fact that the agent is in that state) could be a thing believed to be the case.

thought, one might think not only that these two views are in agreement; one might also think that they are ultimately the same view differently articulated.

But these thoughts are mistaken. There *is* a deep difference between B-Psychologism and the Belief View. Consider once more that, where Non-Psychologism will hold that Florence acts for the reason that roses, B-Psychologism holds that Florence acts for the reason that ROSES *: *I believe* that roses will cheer up my mother. When we keep this in mind, the difference between B-Psychologism and the Belief View becomes clear. The main difference can be seen in the ontological commitments of the two views. B-Psychologism says that Florence's reason is *identical to* or perhaps *constituted by* the fact that she believes that ROSES. But the Belief View makes no such commitment. The Belief View says that Florence's believing that ROSES is a *necessary condition* on her acting for the reason that ROSES. Indeed, the Belief View is compatible with Non-Psychologism: one who holds the Belief View can also hold that the reason for which Florence buys roses is some non-psychologistic entity. But B-Psychologism is, by definition, inconsistent with Non-Psychologism. So B-Psychologism and the Belief View are clearly not the same view, and are possibly at odds.

Can Psychologism Accommodate the Belief View?

But might they be made not to be at odds? Given the plausibility of the Belief View, any Psychologistic theory of reasons should aim to accommodate it if possible. So let us imagine a theorist who, taking it that the Belief View is both distinct from and perhaps equal in plausibility to B-Psychologism, attempts to combine the two theories in building a more general theory of acting for reasons. Call this combination of these two theories **BB-Psychologism**.

What would BB-Psychologism say about Florence? We have already seen that B-Psychologism tells us that Florence buys roses for the reason that ROSES *. The Belief View tells us that in order for Florence to Φ for the reason that ROSES *, Florence must at least believe that ROSES *. So the B-Psychologist, to accommodate the Belief View, must claim that Florence believes *that she believes* that roses will cheer up her mother. There is no logical barrier to such a view — there is no internal inconsistency in the conjunction of these claims. But there is certainly a serious implausibility. To see why, consider a slight alteration of the previous case.

UNREFLECTIVE FLORENCE: Florence’s mother is in a funk. To cheer her up, Florence decides to buy her some flowers. So, she goes to the store to get some flowers. Taking it that roses are her mother’s favorite flowers, she picks out some roses. As she checks out, the cashier asks her, “Why are you buying roses — what’s the occasion?” Florence responds by saying, “I am trying to cheer up my mother, and roses are her favorite flowers. The cashier then asks Florence whether she believes that roses will cheer up her mother. “Obviously,” Florence replies, growing impatient. The cashier then asks Florence whether she believes that she believes that the roses will cheer up her mother. She responds with puzzlement, and says, “I guess I had not really thought about whether I believed that I believed that.”

I wish to make two points about this case. The first point is that the case of UNREFLECTIVE FLORENCE is clearly coherent. The story makes sense, and Florence appears to be a normal — even reasonable — character here. But if BB-Psychologism is true, the story should not make sense; and it certainly could not all be true. For, if BB-

Psychologism is true, Florence *must* believe that she believes that roses will cheer up her mother. Otherwise, she cannot be acting for the reason that she believes that roses will cheer up her mother. But this is a clear *reductio* against BB-Psychologism. It is plainly false that, in order for Florence to be motivated to act as she does, she must believe that she believes that roses will cheer up her mother.

There is a second layer to this point. Florence might, like most of us, not often form second-order beliefs at all — much less *every* time she is motivated to act. What is more, though, there are agents for which it straightforwardly implausible that they *could* have the second-order beliefs that BB-Psychologism requires them to have. Consider two-year-old Thomas, who asks to go to see the dogs in the park.⁷⁶ When his father asks him why they should do that, Thomas says, “It’s fun to see doggies in the park.” Despite his unadvanced age, I think we would naturally take this as a statement of Thomas’s reason for going (or perhaps, wanting to go) to the park. Now, it seems highly unlikely to me that Thomas introspected or inquired at all about his own psychological state concerning the matter — much less that he thereupon formed the belief that he believed that it would be fun to see dogs in the park. Indeed, he might not even have the cognitive resources to approach the issue of his second-order beliefs on the matter. But suppose Thomas’s father asks him whether he believes that he believes that it would be fun to see dogs in the park, and Thomas replies that he does not (or perhaps, more realistically, with a gesture that indicates that he is unsure how to answer). It would be quite a cruel and suspect bit of parenting if Thomas’s father therein found grounds to withhold the trip to the park from Thomas, saying, “Well, it sounds like you have no reason at all to want to go to the park!”

⁷⁶ Perhaps the dogs themselves should be considered here as well. Take Spot, who is digging over there because that is where his bone is buried. Does Spot believe that he believes that that is where his bone is buried? It does not seem likely; indeed, it is not clear that he *could* believe such a thing.

Thomas has clearly stated his reason, and his apparent lack of a second-order belief on the matter does not create any confusion about what that reason is.

The trouble here is not only that BB-Psychologism implausibly commits us to insisting that agents have second-order psychological states where we have no real reason to suspect that they do have them (or where we have reason to suspect that they do *not* have them). The deeper trouble — and my second point here — is that BB-Psychologism implies that it is the having of these second-order states that enables agents to respond to (and be motivated by) the considerations in question. But that is surely implausible. It is not some strange second-order belief that allows Florence to be motivated to cheer up her mother. A first order psychological state, such as believing that roses or wanting to cheer up her mother, seems sufficient for that. This is a point on which the Belief View and B-Psychologism, taken separately, should agree.

Another way of seeing this point is through the lens of action explanation. The reasons for which agents act, common thinking holds, play a crucial role in explaining actions. Recall our very plausible axiom of

REASONS-EXPLANATIONS: a motivating reason is the kind of thing that can play a role in explaining an agent's action.

Given the truth of (something like) REASONS-EXPLANATIONS, any cognitive constraint on acting for a reason effectively places a constraint on explaining actions done for reasons. If we hold the Belief View, for example, we are committed to a view on which we cannot straightforwardly appeal to r in a reasons-explanation of A 's Φ ing without supposing that A believes that r . Similarly, BB-Psychologism is committed to a view on which one cannot appeal to r in a reasons-explanation of A 's Φ ing without supposing that A believes that A

believes that *r*. But cases such as UNREFLECTIVE FLORENCE and two-year-old Thomas show such a view to be incorrect. We can easily explain Florence's buying the roses, for example, by saying that "She is buying the roses for the reason that her mother is in a funk," or that "She is buying the roses because she believes that her mother is in a funk." Offering these explanations does not require supposing that Florence believes that she believes that her mother is in a funk. Furthermore, any view that insists on such a requirement will, I fear, be quite lacking in explanatory ability; for it will provide *no* reasons-explanation (indeed, it might insist that there is no such explanation) of any action for which we cannot suppose that the agent has the requisite second-order psychological states. Why should that be so? If BB-Psychologism is true, then whenever *A* Φ s for a reason, there must be a suitable consideration which *A* believes that *A* believes to be the case. It follows, then, that if there *is* no such consideration at hand, BB-Psychologism is committed, on pain of inconsistency, to saying that it is not the case that *A* Φ s for a reason. By definition, one cannot have a reasons-explanation of *A*'s Φ ing without giving a reason for which *A* Φ s.

Psychologists might be tempted to reply by suggesting that the culprit here is the Belief View, and not B-Psychologism. After all, all I have shown is that the *combination* of the Belief View and B-Psychologism produces implausible results. This leaves open whether the best way to avoid this implausibility is to give up B-Psychologism, or instead to give up the Belief View. Given the plausibility of B-Psychologism, they might say, what we have learned here is that the Belief View cannot be the best theory of the cognitive constraint on acting for a reason.

As I observed earlier in this chapter, the Belief View is not unquestioned. The cognitive constraint is the subject of controversy, and there are many who think that the Belief View does not provide the correct constraint. However, those who reject the Belief

View typically do so in favor of a *stronger* cognitive constraint on acting for a reason.⁷⁷

Recall, for example, that the only established competitor to the Belief View is the

Knowledge View: in order for A to Φ for the reason that r , A must know that r .

Endorsing the Knowledge View will be of no help in resolving the tension observed here. Indeed, it seems to make things even worse. Instead of insisting only that every instance of acting for a reason is accompanied and made possible by a second-order belief, the incorporation of the Knowledge View would force Psychologism to say that acting for a reason is always accompanied and made possible by some second-order *knowledge*; and also that we can give a reasons-explanation of A 's Φ ing for a reason only if we suppose that there is some r such that A knows that A believes that r . This brand of Psychologism would be forced to say, for example, that little Thomas *knows that he believes* that it would be fun to see the dogs in the park — and also that Thomas's so knowing is what enables him to be motivated by the consideration in question. As implausible as the combination of the Belief View and Psychologism is, the combination of the Knowledge View and Psychologism is even more implausible.

It seems, then, that the tension observed here cannot be plausibly resolved by rejecting the Belief View in favor of a stronger view on the cognitive constraint. Might these problematic results be avoided by adopting a *weaker* cognitive constraint? This strategy is no more promising. This is partly because, as I have illustrated in this chapter, there is a strong independent case for the truth of the Belief View. It is plausible that every situation in which A Φ s for the reason that r looks to be a situation in which A must

⁷⁷ The recent work of Daniel Muñoz [2017] might be an exception, though his work concerns acting for a purpose rather than acting for a reason.

(at least) believe that r . Cases like Unbelieving Edmund simply do not make sense. But even if the Belief View turned out to be too strong, this alone would not be enough to resolve the tension here. For, as I said at the outset, there must be *some* positive cognitive constraint on acting for a reason. This is, I think, all that is required for Psychologism to be saddled with similarly problematic results. So long as there is some cognitive constraint on acting for a reason, the combination of that cognitive constraint and Psychologism will always generate a fact of the following sort:

CP-Schema: A [cognitive relation] that A [psychological state] that r .

In the first slot belongs whatever cognitive constraint on acting for a reason one finds most attractive; in the second slot belongs whatever psychological state is centrally featured in the version of Psychologism one finds most attractive. It is simply not clear to me what plausible view of the cognitive constraint could be inserted into this schema such that Psychologism would avoid implausible results of the sort observed here. For, under the CP-Schema, no matter what view of the cognitive constraint or version of Psychologism is under consideration, *any* fact of the form “ A Φ s for the reason that r ” entails that A has some second-order psychological state (suitably related to r) or other. The problem, of course, is that no such entailment holds; the fact that A Φ s for the reason that r clearly does not entail that A has any second-order psychological states at all — much less any *particular* second-order psychological states. So the prospects of resolving the problematic tension by rejecting the Belief View in favor of some weaker cognitive constraint really do not seem any more promising than those of resolving the tension by rejecting the Belief View in favor of some stronger constraint. It looks as though

Psychologism is simply at odds with the existence of *any* cognitive constraint on acting for a reason.

Can Psychologism Do Without the Belief View?

But Here is what I take the upshot to be so far. The combination of Psychologism and any cognitive constraint on acting for a reason generates unpalatable consequences. That being so, one has three choices. One option is to simply swallow these unpalatable consequences, and insist that both acting for a reason and action (reasons-)explanation is always accompanied and made possible by some second-order psychological state. A second option is to give up on the existence of any cognitive constraint on acting for a reason. A third option is to reject Psychologism, and admit the possibility that at least some reasons for acting are things other than psychological states of the agents acting (an option which, it is worth remembering, already has significant independent support).

I hope that the considerations of the previous section have been effective in showing that the first option is quite undesirable. But, lest we give Psychologism short shrift, let us consider the somewhat idiosyncratic second option. Now that we are faced with this choice, the Psychologist might say, perhaps we should revisit the plausibility of the thought that there must a cognitive constraint on acting for a reason. Why should we think this to be a nonnegotiable maxim for theories of reasons? We can begin answering this question recalling the case of

INCONSIDERATE EDMUND: Imagine that you are about to go skating, and you notice Edmund skating on a nearby pond. You notice that when he skates, Edmund keeps to the edge of the pond. As you are about to begin skating, you ask Edmund why he is skating only around the edge of the pond. He responds by

saying, “My reason for keeping to the edge is that the ice in the middle of the pond is thin.” When you thank him for warning you about the thin ice, he says, “Oh — I was not saying that the ice in the middle of the pond is thin. In fact, I had not even considered the thinness of the ice; it had not crossed my mind. Nonetheless, that is my reason for skating near the edge of the pond.”

Whatever one thought of UNBELIEVING EDMUND, the case of INCONSIDERATE EDMUND surely makes no sense. The suggestion that Edmund might have been responding to and acting for the reason that the ice in the middle of the pond is thin, and yet he had not even *considered* — it had not even crossed his mind — whether the ice in the middle of the pond was thin, is incomprehensible. Now, a theory that holds that there is no cognitive constraint on acting for a reason forgoes a principled explanation as to *why* this suggestion is incomprehensible.⁷⁸ Indeed, such a theory has given up the idea that there need be anything strange at all going on in the case of INCONSIDERATE EDMUND. But the case clearly is quite strange, if it is even possible. So we find in the case of INCONSIDERATE EDMUND strong evidence that agents must make *some* cognitive connection to the reasons for which they act.

One need not be convinced that there is a cognitive constraint on acting for a reason by one case. There is also a myriad of theoretical considerations pressuring us in this direction. Consider, for example, that a very common analysis of the notion of a motivating reason for Φ ing is: a consideration taken to favor Φ ing. How can one take r to favor Φ ing without bearing some cognitive connection to r , when *taking to favor* itself seems to be a cognitively-laden relation? Leaving aside the notion of taking to favor: how can

⁷⁸ This might assume that something like *considering that* r is the weakest cognitive (positive) relation one could bear to r . This might not be the case, but it would not be difficult to draw up an analogous case for whatever turns out to be the weakest cognitive relation.

one be motivated to act by r without making some cognitive connection to r ? If motivating reasons are to be anything other than mere causes, this seems a difficult feat to say the least. Indeed, even if motivating reasons *are* causes, it still seems that by far the most likely story as to how these causes do their motivational work is that they do it by bringing about some cognitive event. Consider further the question of what makes something *my* reason. If the ice in the middle of the pond is thin, how does this get to be my reason for acting, rather than *your* reason — or even just *a* reason? The answer that suggests itself is that some cognitive connection I made to the reason (such as believing it to be so, or taking it to favor skating near the edge of the pond) is what makes it my reason. Consider further still the issue of action explanation. Recall the very plausible idea of REASONS-EXPLANATIONS. There is a certain sense of the question, “Why is Florence buying flowers?” which asks for her reason for acting. How can we completely and satisfactorily answer such questions without saying *something* about what is going on with the agent cognitively? The list goes on. It is not just an analysis of a few cases that should lead us to believe there is a cognitive constraint on acting for a reason. There is a whole host of interesting and important philosophical questions concerning actions and reasons which begin to approach being unanswerable if there is no cognitive constraint.

So taking the second option here is not easily done. It seems, generally, that any theory of acting for a reason that is lacking a cognitive constraint will be far less plausible than those that are not so lacking. Unless there is some account as to how Psychologism might be able to offset or mitigate these theoretical costs, perhaps we should give up Psychologism (rather than the cognitive constraint). It is worth asking, then: can Psychologism offset the costs? We saw in the first section that Psychologism shares some important similarities with the Belief View. Of course, we also saw that there is a deep difference between these two views. But perhaps the Psychologist, in light of the

similarities between B-Psychologism and the Belief View, might try to persuade us that a view like B-Psychologism need not suffer any theoretical costs not suffered by the Belief View. Every side agrees, the Psychologist might say, that in the case of FLORENCE'S FLOWERS, any good theory must centrally feature the consideration that ROSES, the claim that Florence believes that roses, and be able to explain Florence's buying the roses by casting it as a response to the consideration that roses. Addressing some of the theoretical pressures just discussed, the Psychologist might say that B-Psychologism, like the Belief View, holds that a belief with a certain content is necessary for rational motivation to take place. It can also provide an account similar to that of the Belief View as to how agents' reasons get to be *their* reasons: this comes about by some cognitive event such as, for example, believing that roses. Analogous points may be made about action explanation. B-Psychologism can accord with the maxim of REASONS-EXPLANATIONS, for example, since in giving any explanation that cites a motivating reason, a Psychologistic theory (which holds that motivating reasons *are* cognitive states) will therein be telling us something about what is going on with the agent cognitively. And so on. Of course, none of this will persuade someone who simply finds cases like INCONSIDERATE EDMUND to be clean counterexamples to any theory without a cognitive constraint. But if the previously mentioned theoretical costs were really what was persuading us away from Psychologism, perhaps we moved to quickly; for Psychologism may be able to give an equally satisfying account on these fronts without adding any cognitive constraint to its theory. Granted, Psychologism will accomplish all of this by saying that Florence is acting for the reason that roses*, while the Belief View can accomplish things via a simpler Non-Psychologistic view, holding that Florence is acting for the reason that roses. But in light of all of the similarities just mentioned, we can see that these two claims really amount to the same

account of what Florence is doing; and suggesting otherwise is just word-smithing and hair-splitting over ontological minutia.⁷⁹

This is not a completely implausible line to take. I think there are two ways of construing this line of resistance; I shall consider them both. One way of seeing this line is as a sort of last-ditch effort to uphold the bold claim that there is no cognitive constraint on acting for a reason. Cases like INCONSIDERATE EDMUND like aside, the line goes, nothing of real importance is lost by (for example) endorsing B-Psychologism and rejecting the Belief View. On this way of taking things, though, this line of resistance can easily be seen off as misguided. For it does not address the real kernel of wisdom revealed in cases like INCONSIDERATE EDMUND. The wisdom is that it is incomprehensible that *A* might Φ for the reason that *r* and yet not so much as consider the idea that *r*. The point here is structural; it makes no difference whether the content of the reason is that *p*, or instead (as Psychologism insists) that *A* believes that *p*. Furthermore, the theoretical pressures here are only *apparently* relieved in the absence of a genuine cognitive constraint. Consider again, for example, the common analysis of what a motivating reason is. Even if Florence's reason is that *she believes* that roses will cheer up her mother, the question remains how she comes to take *that* consideration to favor her buying flowers without so much as considering it. The same question arises vis-à-vis the earlier question about motivation. What of the individuation and possession of reasons? Even if Florence acts for the reason that she believes that roses will cheer up her mother, how does *that* consideration get to be *her* reason for acting? Saying that it happens via her believing that roses will cheer up her mother is vacuous. If Florence's reason is *that she believes that roses*, then saying that this consideration comes to be Florence's reason (or that Florence comes

⁷⁹ Thanks to David Sosa for helping me to see this point of view.

to have this reason) by her believing that roses amounts to saying that the reason gets to be hers by merely being the case. But what we were wondering in the first place was, as it were, how some reason goes from merely obtaining to being a reason that an agent *has*. According to the view on offer, though, Florence need not have any other cognitive state in order to act for the reason in question. Finally, there was the issue of action explanation. It is true enough that Psychologism has no trouble explaining actions. It can surely, for example, offer explanations of the form, “ A Φ d because A believed that p .” But given that it seems to be a bit of a mystery, on this view, how A comes to act for the reason that A believes that p , it is not clear how confident we should be that this brand of Psychologism can even offer genuine reasons-explanations — explanations of the sort that would satisfy the maxim of REASONS-EXPLANATIONS.

So I think, out of charity, we should not take this line of resistance as a wholesale rejection of the cognitive constraint. Indeed, I think the foregoing show that all parties to the debate should concede that any plausible theory of acting for a reason needs something to at least fill the theoretical role of a cognitive constraint. Here, then, might be the second (and more charitable) way of seeing the current line of resistance: as trying to persuade us that Psychologism, in a way, itself offers a cognitive constraint on acting for a reason. While the Belief View provides a cognitive constraint by making A 's belief that r a *necessary* condition on A 's Φ ing for the reason that r , however, Psychologism does so by making A 's belief that r a *constitutive* condition on A 's Φ ing for the reason that r . Seeing things through this lens, one might think that there is ultimately not much to choose between the two theories in question.

I think, however, that there is a clear choice between these two theories. Let us suppose that the Psychologist has convinced us that a view like B-Psychologism is just as plausible as the Belief View. The resulting view still suffers a serious, and in my view

unacceptable, cost. What we are supposed to be persuaded of here is that the theoretical role of the Belief View can be filled by a view like B-Psychologism. The need for a cognitive constraint, this way of thinking goes, is satisfied by Psychologism. But this is tantamount to claiming that the notion that whenever A Φ s for the reason that r , r is something A believes (namely, p), can be reduced to the notion that whenever A Φ s for the reason that r , r is the fact that A believes that p . This brand of Psychologism tells us that endorsing this latter claim is (*inter alia*) really a way of endorsing the former claim. As such, the resulting theory is one that has no room for a legitimate distinction between A 's Φ ing for the reason that r and A 's Φ ing for the reason that A believes that r . But there clearly is a distinction here, as Jonathan Dancy has so compellingly observed:

Someone who believes that there are pink rats living in his shoes may take that he believes this as a reason to go to the doctor or perhaps a psychoanalyst. This is quite different from the person who takes (his belief) that there are pink rats living in his shoes as a reason to call the pest control officer. Such contrasts show that we will distort what we have been calling the light in which the agent acted, or the agent's reasons, if we insist that they are [all] properly specified as 'that he believed that p .'⁸⁰

This example illustrates the clear and important difference between Φ ing for the reason that r and Φ ing for the reason that one believes that r . In most cases, the difference can be located in what these two considerations are taken to favor. There being pink rats in my shoes seems to favor calling the pest control officers; my *believing* that there are pink rats in my shoes seems to favor calling my psychiatrist (or taking a moment to rub my eyes, collect my thoughts, or what have you). Any theory of acting for a reason that cannot countenance these facts, or the distinction undergirding them, should be thrown out.

⁸⁰ See Dancy [2000: 125]; and also Alvarez [2010: 48-49], Anscombe [1989: 381], Hyman [1999: 444-446], and Joseph Raz [1999] for corroborating treatments of similar examples. I am not here implying, as many do, that such examples show directly that Psychologism is false. Instead, I claim only that such cases illustrate an obvious distinction.

It is worth noting here, as it was in the previous section, that these problems are not confined to theories that endorse the Belief View (rather than some other cognitive constraint). For the general underlying problem here is a kind of conflation of the idea that agents must bear some significant cognitive relation to any consideration in order to be motivated by it with the idea that the considerations which motivate agents are their cognitive or psychological states. Any theory that attempts to reduce a cognitive constraint on acting for a reason to a psychologistic ontology of reasons will be guilty of such a conflation. This might not be grounds for rejecting Psychologism if *every* plausible theory of reasons were bound to make this conflation. But a theory that accepts a cognitive constraint on acting for a reason (such as the Belief View or the Knowledge View), while rejecting Psychologism, need not be guilty of such conflation. This is easily shown by considering what such a view would say about Dancy's case involving the pink rats. If A 's reason is that *there are pink rats in my shoes*, the Belief View says that A can act for that reason (and call, for example, the pest control officer) so long as A believes it to be so. If A 's reason really is that *I believe that there are pink rats in my shoes*, the Belief View says that A can act for that reason as well (and call, for example, the psychiatrist) — so long as A believes *that* to be so. Such a view clearly has no trouble respecting the distinction between placing a cognitive constraint on acting for a reason and conceiving of a reason for acting *as* a cognitive state. Now, one might reasonably think that there is not much to choose, all else equal, between the view that

Whenever A Φ s for the reason that r , r is the fact that A believes that p ;

and the view that

Whenever $A \Phi$ s for the reason that r , r is something A believes (namely, p).

But there is surely an easy choice between a theory that recognizes the difference between these two views and a theory that does not.

Summary Remarks

Where does all of this get us? We started out by observing that there are two seemingly plausible views in the theory of reasons: the Belief View and B-Psychologism. We then observed that when we try to combine these two views, the resulting view is deeply implausible. That being the case, we were left with three choices. The first choice is to simply endorse the resulting view despite its serious shortcomings. The second choice was to give up the idea that there is any cognitive constraint on acting for a reason — that when $A \Phi$ s for the reason that r , A must bear some cognitive connection to r . The third choice was to give up Psychologism, and admit that motivating reasons for acting might be things other than the psychological states of agents.

I think that the picture of acting for a reason discussed in subsection one shows that the first option is to be avoided if at all possible. It is simply not plausible that acting for a reason or action explanation is always accompanied and made possible by second-order psychological states. The second option is shown to be implausible, at least initially, by cases like Inconsiderate Edmund. It is simply not clear how such cases could be possible. We are also pressured away from the second option in view of how difficult questions about motivation, the individuation and possession of reasons, and action explanation become without any cognitive constraint. Though Psychologism may not be able to counter the hard intuition that cases like Inconsiderate Edmund are not possible, it might be able to go some way toward satisfying concerns about these theoretical

pressures. In doing so, however, Psychologism must effectively collapse a distinction that any plausible theory should be able to countenance. The conflicting Belief View, on the other hand, can preserve the distinction. I take this to show, all together, that the second option is also not the most preferable. The best way to resolve the tension between these two views, I conclude, is to give up Psychologism, and preserve the commitment to the cognitive constraint on acting for a reason (whatever the correct view of that may be).

Chapter Three: Arguments from Error

At this point, I hope to have made the case for two main thoughts, which are answers to central questions in theorizing about action and reasons. The first is that in order for an agent to act for a reason r , it need not be the case that r . The second is that in order for an agent to act for the reason that r , the agent need only believe that r . I have also argued that a byproduct of this second thought is that it is not plausible that whenever an agent acts for the reason that r , r is a psychological state of (or fact about) the acting agent. For each of these two main claims I have tried to present the intuitive case, and then defend them from what I take to be the best objections. In this chapter I want to build on the work of the previous two chapters. I think that we can extract an important lesson from the discussion surrounding these two issues (the factivity of acting for a reason, and the cognitive constraint on acting for a reason). This lesson is one that can give us insights as to the nature of action explanation, such that we can construct better overall theories of acting for a reason.

As may be apparent by now, it is a common thought among philosophers that theorizing about a certain class of cases, which have been called “error cases”, is crucial to settling some important and central debates. These cases are those, roughly, in which the agent errs in some way in arriving at a view about the way the world is, and then acts in the light of that view of the world. Theorists divide on how they respond to error cases. But despite their differences, many theorists claim to glean motivation for their views from these cases; and many do so in the same way. They claim that while their theory has no trouble giving good reasons-explanations of actions in such cases, other theories are explanatorily deficient.

My primary purpose in this chapter is to show that this kind of argument is flawed. Interestingly, the most problematic occurrences of this strategy are seen within arguments that have played a large role in gaining certain views the popularity they currently enjoy. I will examine several of these arguments and demonstrate that they fail; I will then demonstrate that we can abstract away from the particulars of each argument to see that this *kind* of argument is flawed, and so each instance of it is bound to fail. The fatal flaw is that there is quite a large gap between what can be said about error cases themselves and the quite general claims about acting for reasons that theorists claim the cases establish; and this gap is one that can only be bridged by unmotivated and implausible claims about the nature of action explanation.

The course of the chapter is as follows. In the first section, I review and present a general sketch of the kind of reasoning that is supposed to take us from error cases to some well-received views in action theory. I separate error cases into two kinds, which I will call “factual” error cases and “procedural” error cases. In the second section I treat arguments based on factual error cases; in the third section I treat arguments based on procedural error cases. In each of these two sections, I demonstrate that in each instance of the argument from error, the most crucial moves are not warranted — at least, not without the aid of some auxiliary constraint on action explanation. Problematically, the needed constraints are rarely made explicit by advocates of the arguments in question; but I show that once they are made explicit, it is clear that they should be rejected. In the fourth section, I return to the project I postponed at the end of the first chapter, and try to lay out a fuller picture of action explanation compatible with a packaging of the Nonfactive View_M, the Belief View, and Non-Psychologism. I conclude in the fifth section with some remarks about where all of this might leave theorizing about acting for a

reason. Since one might take analyses of error cases to be the deciding factor for these debates, I briefly inquire as to where else we might look.

In making these points I hope to encourage some reevaluation of the current landscape in action theory — but I also hope for more than that. Clearly, the points made in this chapter are directly applicable to theorizing about the general nature of explanation, and I hope to stimulate further thought on that front. Furthermore, given the ease with which claims about reasons are transplanted from one normative domain to the next, I hope that the points made here also provoke thought in epistemology and ethics — about what it is to believe for a reason, about the nature of evidence, about responding to moral reasons, and more. I say more on that front in the fourth chapter.

THE USE OF ERROR CASES IN ACTION THEORY

As I said, much of these current debates revolves around accounting for what can be called **error cases**: cases wherein the agent errs in arriving at a view about how the world is, and then acts in the light of that erroneous view of the world. As we will soon see, the debates around error cases actually involve two different kinds of error cases. But what I want to illustrate in this first section is how error cases are generally used to pressure theorists away from otherwise intuitive and appealing views. Supposing that (as I have argued) the intuitive package of views combines the Nonfactive View_M, Non-Psychologism, and the Belief View, one way of seeing things is that there is a train of thought involving error cases which pushes us to abandon each of these views in succession.

The train of thought begins with the more common of the two types of error cases, which I will call **factual error cases**. These are cases in which the agent comes to an erroneous view about the way the world is, and acts in the light of that erroneous

view of the world. There is a plain variation of our case of Edmund that fits this description. We can imagine that Edmund has gone out skating, taking it that the ice in the middle of the pond is thin. So, when he skates, Edmund keeps to the edge of the pond. In fact, the ice in the middle of the pond is not thin. In such a case, Edmund is simply wrong about the way the world is; he has made an error in taking it that thin ice: the ice in the middle of the pond is thin. That notwithstanding, we have also seen that it is natural — even plausible — to accord with the Nonfactive View_M and take the case as one in which Edmund acts for the reason that thin ice. However, we have also seen that there is an influential argument against this natural reaction. For convenience, here is the earlier passage from Jonathan Dancy explaining how factual error cases are supposed to cause trouble.

This Edmund example faces us with an inconsistent triad:

1. We explain an action by giving the reason for which it was done.
2. Edmund's reason for keeping to the edge was that... the ice in the middle was thin
3. That the ice in the middle was thin is not (any part of) what explains Edmund's keeping to the edge.

This is now recognizable as one variation of the Argument from the Factivity of Explanation (AFE) that was addressed in the first chapter. These three claims, the argument goes, are jointly inconsistent. Since the first and third claims are articulations of the axiomatic reasons-explanations and factive explanations, we are forced to give up the otherwise plausible Nonfactive View_M in favor of some kind of Factive View_M .

As I have suggested, the argument, in a nutshell, comes to the claim that the Nonfactive View_M is incompatible with giving satisfying reasons-explanations in factual error cases. Notice, though, that this argument (even if successful) only tells us what Edmund's reason could *not* have been; it does not tell us what Edmund's reason is or must have been. But we do not want our theorizing to stop at just robbing Edmund of (what he

and we took to be) his reason for acting. After all, he does not seem to be doing anything very strange, and we do not want our theory to make him out to be unintelligible. So we should replace thin ice with some other reason — ideally, one that would help us to make sense of what Edmund is doing. This theoretical gap is thought to give an opening to Psychologism. How so? Again for convenience, here is the earlier passage from James Lenman explaining how such cases might be used to motivate Psychologism.

It is after all extremely natural to explain the agent's fleeing the building by reference not to his belief that there was a fire but to the fact that there was a fire. Or, when we ask, Why does Angus punch his boss? we naturally reply, Because he has been fired; not, Because he believes he has been fired. It's the fact to which the belief answers, not the belief itself, to which the explanation of such actions naturally refers...

The biggest headache for anti- psychologists such as Dancy however is furnished by cases where the agent's belief is *false*. The fact of Angus' being fired is naturally adduced to explain his punching his boss in cases where he has indeed been fired. But in cases where Angus punches his boss, believing mistakenly that he has been fired, it seems quite wrong to say he so acts because he has been fired. In such a case we surely must retreat to a psychologised explanation if we are to have a credible motivating reason explanation at all.

The core of the argument here is quite similar to the previous argument for the Factive View_M. It assumes that reasons for acting play an explanatory role, and that action explanation is factive. On those grounds, the argument goes, we must dismiss the idea that a motivating reason could be something that is not the case. But what then could it be? What feature of these cases could accord with these axioms and plausibly be considered a motivating reason? It might seem that the only good candidate is some fact about the agent's psychology. Instead of saying that Angus punches his boss for the reason that fired: he has been fired, we should say that Angus punches his boss for the reason that fired*: *he believes* he has been fired.

So one can see how factual error cases might lead one to abandon the intuitive Nonfactive View_M and Non-Psychologism in favor of some Psychologistic Factive View_M. Furthermore, it is worth noting, if my argument at the end of the previous chapter is correct, the truth of Psychologism all but rules out the truth of the Belief View.

I said a moment ago that there are two different kinds of error cases. We have so far only looked at one kind; but it is the second kind that is often instrumental in moving from Psychologism to a view often called

Disjunctivism: If A Φ s for the reason that r , then:

either A knows that r , and so Φ s for the reason that r ;

or A Φ s for the reason that A believes that r .⁸¹

Disjunctivism, in a sense, incorporates the wisdom of Psychologism. But it is perhaps more flexible than Psychologism. Like Psychologists, Disjunctivists are impressed by the arguments involving factual error cases. They think that Angus' reason, for example, could not have been that he was fired. They also think that the best explanation of what he is doing is a psychologized one, such as the claim that he believes that he was fired. But while Psychologism holds *all* reasons to be psychological facts about agents, Disjunctivism appears to allow that some reasons may be located outside the agent. Pairs of cases are often illustrative here. Compare Angus with Agnes, who has been fired. Agnes knows that her boss has fired her; so, when she is packing up her things to leave the office, she punches her boss. Psychologism holds that both Angus and Agnes act for the reason that

⁸¹ Here I have given what I hope to be a fair rendering of Hornsby's [2008: 252] account. Her account uses "because" where I have "for the reason;" but she explicitly states that her "because" should be understood in that way. As Dancy [2008] demonstrates, it takes a bit of work to formulate the account as both disjunctive and as about acting for reasons. Admittedly, the view is not shared in exactly this form by all who identify as "Disjunctivist". For example, Hyman [2011] distinguishes between being guided by the fact that r and treating r as a premise (which does not count as acting for a reason); McDowell [2013] prefers to talk of acting in the light of a fact.

fired*. Disjunctivism, contrastingly, can say that Agnes acts for the reason that fired while Angus acts for the reason that fired*. The trigger for the alternative (psychologized) gloss, of course, is lack of knowledge on the part of the agent.

Notice that this way of advocating for Disjunctivism over Psychologism must include an argument for what I have been calling the Knowledge View. As we saw, there is a debate between those who endorse the Knowledge View and those who endorse the Belief View. Like Psychologists and Factivists, adherents of the Knowledge View think that error cases give their view a decisive advantage over their opponents. But adherents of the Knowledge View are interested in a different kind of error case than Psychologists and Factivists. Here again is Jennifer Hornsby, laying out the kind of argument that has won the Knowledge View an increasing amount of favor:

Edmund... believes that the ice in the middle of the pond is dangerously thin, having been told so by a normally reliable friend. But Edmund's friend didn't want Edmund to skate in the middle of the pond (never mind why), so that he had told Edmund that the ice there was thin despite having no view about whether or not it actually was thin. Edmund, then, did not keep to the edge *because* the ice in the middle of the pond was thin... The fact that ice was thin does not explain Edmund's acting, even though Edmund did believe that it was thin, and even though the fact that it was thin actually was a reason for him to stay at the edge...

Edmund here is used to show that someone's having a true belief (even a justified one) which explains their acting is not sufficient for them to have acted for an (F)-type reason. We saw that an (F)-type reason has to have registered with an agent if they are to have acted on it. We see now what this amounts to. A condition of Φ -ing for the reason that p , when one believes that p , is that one knows that p .

Notice This reworked example of Edmund involves a different kind of error from the ones we encountered in the arguments over factivity and Psychologism. Edmund's error (if it can be called that) is not a pure factual error — indeed, Edmund has a *true* belief about whether the ice in the middle of the pond is thin. Rather, this is what I will call a **procedural error case**: a case wherein the agent in some way errs in the process of

coming to a view of the world, and then proceeds to act in the light of that view of the world. Edmund has arrived at his belief that thin ice in such a way that his belief does not amount to knowledge. In that sense Edmund's belief-forming process is erroneous, and he is acting in the light of an erroneous view of the world.

As we have seen, advocates of the Knowledge View hold that this makes it impossible for Edmund to act for the reason that thin ice. The style of argument here is often similar to that of the argument for the Factive View and for Psychologism. Advocates of the Knowledge View observe that reasons for acting play an explanatory role, and that explanation is factive. On the one hand, the Disjunctivist is impressed by the arguments for the Factive View and Psychologism *in factual error cases*. But what about cases in which the agent makes no factual error? One might think there is little motivation to psychologize reasons in such cases, since there will be no tension between the falsity of the agent's reason and the temptation to use the reason to explain the agent's action. But procedural error cases are supposed to show us that, at least outside of factual error cases, the agent must know her reason for acting; otherwise we cannot give a reasons-explanation of her action. Disjunctivism, its advocates claim, is the only view that captures both of these truths.

In this section I have revisited and given a rough sketch of three influential arguments for popular views about acting for reasons. I have also illustrated how theorists might progress in their analyses of error cases, and shown how these influential arguments might be connected. All three of Psychologism, Disjunctivism, and the Knowledge View might rely on the argument in favor of the Factive View_M; Disjunctivism further relies on the arguments for Psychologism and the Knowledge View.⁸² Although these views are

⁸² I think that what I say here is true for a standard sort of Disjunctivism. Two possible exceptions worth noting are Alvarez [2010] and McDowell [2013], who might both be counted as Disjunctivist in some sense. But both Alvarez and McDowell refuse arguments for Psychologism, and Alvarez herself does not

different (even occasionally at odds), the main arguments for these views share a common core: they all claim that error cases pose an explanatory challenge that only one view can face up to. The challenge, specifically, is to reserve a significant explanatory role for reasons in error cases. All competing views, theorists who employ these arguments claim, are explanatorily deficient in error cases. But is that true?

ARGUMENTS FROM FACTUAL ERROR

In this section, I will continue my argument that it is not true — or at least, that such claims are not earned by the lines of reasoning just laid out. I say that I will continue (rather than begin) my argument because some of the necessary work was completed in the first two chapters. What I hope to show out now, drawing on the work from the previous two chapters, is this. All arguments from error have in common a two-step structure. The first step in such arguments is what I will call **reason-robbing**: the argument is first supposed to establish what the agent's reason could *not* have been. We are supposed to be able to infer on this basis that certain views of acting for a reason are false. The second step in arguments from error is what I will call **reason-replacing**: the argument is then supposed to establish what the agent's reason *must* have been (if not the thing of which the agent was just robbed). In what follows, I will argue that both crucial steps rely on implausible constraints on action explanation. I will focus my critique on just the first step, and demonstrate that arguments from error do not even get so far as to warrant the reason-robbing that they purport to do.

There are some who have already criticized the use of similar arguments in action theory. Maria Alvarez says that

explicitly endorse the standard arguments for the Knowledge View. Furthermore, as we will see, it may be possible to understand McDowell's argument for the Knowledge View as not relying on any argument for the Factive View_M.

It can be argued that reasoning from error cases to a general conclusion about what motivating reasons are is a form of the ‘argument from error’ and that argument is, at best, unreliable. For, as has been pointed out, if critical examination of cases of error or failure... leads to some conclusion regarding what to say about *those* cases, it does not follow that we must say the same about the *veridical* or *success* cases.⁸³

I agree with Alvarez that even if an argument from error showed (for example) that Psychologism gives the correct verdict *in error cases*, it would not follow that Psychologism is generally true. But I want to argue that arguments from error do not even give us the correct diagnoses *in error cases*; and I want to do so in what I take to be a novel way.

Since factual error cases are more common to the discussion of the major issues in the area, I will start with those. As a first step, here is a rough pass at laying out the the argument from factual error cases for both the Factive View_M and Psychologism.

- [1] Whenever *A* Φ s for the reason that *r*, we can explain *A*’s Φ ing by giving *r*.
- [2] If it is not the case that *r*, then *r* cannot be any part of what explains anything.
- [3] So, whenever *A* Φ s for the reason that *r*, *r* cannot be something that is not the case.
- [4] In some cases, *A* appears to Φ for the reason that *r*, though it is not the case that *r*.
- [5] In such cases, *A* at least believes that *r*.
- [6] The fact that *A* believes that *r* can explain *A*’s Φ ing.
- [7] In such cases, there is no other fact that can explain *A*’s Φ ing.
- [8] So, in such cases, the fact that *A* believes that *r* is *A*’s reason for Φ ing.

⁸³ See Alvarez [2010: 135-136]. See also Dancy [1995: 423-428].

Here claims [1]-[3] are an attempt at representing the influential argument for the Factive View_M (laid out earlier from Jonathan Dancy). They are supposed to accomplish the reason-robbing step of teaching us that Edmund's reason for acting could not have been anything that is not the case. Claims [4]-[8] are an attempt at representing the influential argument for Psychologism (laid out earlier from James Lenman). They are supposed to accomplish the reason-replacing step of teaching us that Edmund's reason for acting must have been a psychological fact about him.

The progression from [1] to [3] provides a nice starting point, because the inference there represented is in danger of being fallacious. It also includes a claim — claim [2] — which we have seen no theorist really accepts. Our discussion in the first chapter equips us to see what needs to happen here: either [1] or [2], or possibly both, needs to be revised. There is a logical gap created by the phrases “we can explain A 's Φ ing by giving r ” in [1] and “ r cannot be any part of what explains anything” in [2]. I suggested in the first chapter that this style of argument has nonetheless managed to persuade many who have not noticed the subtle ambiguities (and here, equivocation) in the notion of *being explanatory*. In illustrating the intuitive appeal of certain nonfactive and non-psychologistic views, for example, Lenman expresses the axiom of REASONS-EXPLANATIONS by saying that we “explain... by reference” to the reason, that the “explanation... naturally refers” to the reason, and that the reason is “naturally adduced to explain” the action in question. In noting the supposed problem that the case of Angus poses for certain views, though, Lenman expresses the axiom by saying that “it seems quite wrong to say he so acts because he has been fired”. The first, intuitive way of capturing the axiom is perhaps adequately expressed by one or more of our disambiguations from chapter one. Consider, for example,

[1a] A reason is the kind of thing that can be given in an answer to a “Why?” question.

It also seems consistent with the explanatory role I have suggested the Nonfactive View_M could reserve for false motivating reasons: such reasons are legitimately explanatory in that the action can be explained by specifying them.

But Lenman’s second way of stating the axiom would be more precisely expressed by some other disambiguation. Plausible candidates might include the following.

[1a*] A reason is the kind of thing that can be given *as* an answer to a “Why?” question.

[1b] A reason is the kind of thing that can replace *q* in a claim of the form “*P* because *q*”.

[1c] A reason is the kind of thing that can be the *explanans* of an action.

Now, as saw in the first chapter, this second group of claims is importantly different from, and possibly at odds with, claim [1a]. Furthermore, for claim [1a] to be used in a reason-robbing move, we also saw, it would need to be accompanied by the claim that

[2a] Anything that can be given in an answer to a “Why?” question must be a fact.

This is the variation of FACTIVE EXPLANATIONS that must be paired with [1a] to logically rule out the possibility of false motivating reasons. But claim [2a] is deeply implausible (and endorsed by no one, so far as I know). As such, it presents an uncontroversial

example of the general kind of lesson I want to draw about arguments from error. Why, exactly, is claim [2a] implausible, and not endorsed by anyone? It is so because it places an incredibly implausible constraint on action explanation, which we can call

NO FALSE THAT-CLAUSES: if some claim q explains some other claim p , then q cannot contain any false that-clauses.

This claim is clearly implausible, and should be rejected by everyone. Though FACTIVE EXPLANATIONS may be true, NO FALSE THAT-CLAUSES clearly does not capture the sense in which this is so. Psychologized explanations illustrate why. If the NO FALSE THAT-CLAUSES constraint holds, then in error cases we cannot explain the agents' actions via their false beliefs. But such psychologized explanations are extremely promising. More importantly, as the passage from Lenman pointed out, factual error cases are the very cases for which psychologized explanations are most appealing. Indeed, the most common move for those who wish to rob Edmund of his reason that THIN ICE is to then replace it with the reason that THIN ICE*: he believes that the ice in the middle of the pond is thin. If we accord with REASONS-EXPLANATIONS (thinking that explanatory claims must at least include the contents of the agent's reason for acting), then this move will lead to psychologized explanations like the one just given. But in factual error cases, then, reasons-explanations are bound to include a falsehood, and therefore violate the NO FALSE THAT-CLAUSES constraint. The lesson here is not that we should disregard such reasons-explanations as unsuccessful. Instead, what we can take from all of this is that any argument from error in which the reason-robbing strategy is held up by NO FALSE THAT-CLAUSES is shown to be quite unpromising.

Much of this has been by way of review, as it were. But I hope now to be moving from specific points about particular arguments to general lessons about theorizing in the area. Error cases *by themselves* present no threat to intuitive views such as the Nonfactive View_M. Even when we know that Edmund is being motivated by a false consideration, that fact alone might not persuade us that that consideration could not be his motivating reason. What I am now calling arguments from error are supposed to take us from the data of error cases to certain views about acting for a reason. Now, we saw in the first chapter that some such arguments trade on implausible premises, such as [2a]. But what I hope to reveal now is that these implausible premises also imply, entail, or simply amount to implausible views about the nature of and rules governing either action explanation or explanation in general — views such as NO FALSE THAT-CLAUSES. My claim, then, is that arguments from error in general depend on implausible views about explanation; without these views, it is quite difficult to develop arguments from error cases which threaten the intuitive positions of the Nonfactive View_M, Non-Psychologism, or the Belief View. The NO FALSE THAT-CLAUSES constraint presents an easy target as far as all of this goes. I now turn my attention to more difficult targets.

The Explanatory and Explanantia

The second grouping of claims above points toward one such target. For convenience, here is that group of claims once more.

[1a*] A reason is the kind of thing that can be given *as* an answer to a “Why?” question.

[1b] A reason is the kind of thing that can replace *q* in a claim of the form “*P* because *q*”.

[1c] A reason is the kind of thing that can be the *explanans* of an action.

We saw in the first chapter that various arguments from error may rely on one or more of these claims. For example, we saw that someone might try to argue for the Factive View_M as follows.

[1c] A reason is the kind of thing that can be the *explanans* of an action.

[2c] Anything that can be the *explanans* of an action is a fact.

[3c] Therefore, all reasons are facts.

Such an argument would purport to underpin a reason-robbing move in any case where the agent appears to act for the reason that r but it is not the case that r . It might also gesture toward a reason replacing move: whatever else r is, it must be a fact.

Now, [1a*], [1b], and [1c] perhaps imply different ontologies about what it is to be an *explanans* — a thing that explains. That notwithstanding, they seem to share a common thread: when an agent acts for a reason, the reason *itself* is *what explains* the agent's action. This thought is explicit in [1c]. Under the framework suggested by [1a*], one might hold that the *explanans* is the answer to the relevant “Why?” question. Under the framework suggested by [1b], one might hold that the *explanans* is whatever can be substituted for q . Each of these points toward the more general thought that

EXPLANANTIA ARE REASONS: in any reasons-explanation for which the *explanandum* is A 's Φ ing, the *explanans* must be A 's reason for Φ ing (r).

We can imagine how a principle like this might underly some reason-robbing moves, like that implied in [1c]-[3c]. After all, if the *explanans* of any action done for a reason *just is* the reason, then the reason for acting can never be a falsehood. For then we strictly violate our commitment to FACTIVE EXPLANATIONS, which plainly states that nothing that is not the case can explain anything. Here is Jonathan Dancy, illustrating (perhaps unintentionally) how parties on multiple sides of the debate have relied on something like this principle.

Sentences of the form ‘his reason for V-ing was that *p*’ can be true even when it is not the case that *p*... This led me to think that these explanations of action are not factive, since ‘his reason for V-ing was that *p*’ does not entail *p*. And this raised the bewildering question how something that is not the case can be what explains something that is the case. Some thinkers are so put off by this question that they resort to different accounts of the agent’s reason for any case where the agent has made a mistake. The most popular account is [Psychologism].⁸⁴

We can see here how something like EXPLANANTIA ARE REASONS can underpin both reason-robbing arguments *and* responses to those arguments.

The trouble, though, is that it is not plausible that EXPLANANTIA ARE REASONS.

Here again is Jonathan Dancy, revealing some of the insight as to why.

If we stick to this thought [that we explain an action by specifying the reason for which it was done], the only possible resolution is that the explanans is not identical with the reason for which the action was done... we can say that what explains the action is that it was done for the reason that *p*, without committing ourselves to saying that what explains the action is that *p*.⁸⁵

In Dancy’s framework for action explanation, the sentences of the form “*A*’s reason for Φ ing is that *r*” are reasons-explanations of actions in the sense that these sentences are *what explain* the actions. As I indicated in chapter one, it seems wholly undeniable to me that sentences of this form present at least one kind of action explanation. I am also

⁸⁴ See Dancy [2014: 89].

⁸⁵ See Dancy [2014: 90].

inclined to agree with Dancy that such explanations work without entailing that r is the *explanans* of A 's Φ ing.

So I think that Dancy's remarks nicely show one way in which EXPLANANTIA ARE REASONS might simply be false. More important, though, is what this principle implies about the general nature of explanation. This principle implies that the explanatory role of something (a reason, for example) consists in being the *explanans* of some *explanandum*. But this deeper constraint on explanation seems to me deeply implausible. In view of the fact that the Nonfactivist may be better off not saying that motivating reasons are *explanantia*, Dancy seems to shy away from saying that reasons *themselves* do any significant explanatory work. He says that, in doing this, he has "yielded to a barrage of criticism" from many, including Wayne Davis.⁸⁶ As Davis puts it,

We need to distinguish the claim that *actions can be explained by reference to reasons* from the claim that *reasons are what explain the action*. The former is true, the latter is false. The statement that my reason for saving was that my son will need money for college does explain why I saved. But it does not follow, and it is not true, that my reason explains my action... my action would have the same explanation even if I were wrong in thinking that my son will be going to college. The claim that motivating reasons are *things that explain actions* is undermined by the fact that motivating reasons are intentional objects. To think of reasons themselves as explanatory is to treat instances of "that p" as referential terms.

I think that the distinction marked by Davis is a genuine and important one. I also think, however, that Davis' distinction admits of another distinction: that between explaining the action and *being what explains* the action. I want to suggest that, even when they are not *explanantia* (again, conceiving of an *explanans* in such ways as *the answer* to the question "Why?", *the proposition* that substitutes in for q in " P because q ;" *the whole set* of sentences that accounts for the *explanandum*; or some other such), there is a legitimate sense in which reasons explain actions.

⁸⁶ See Dancy [2014: 90] and David [2003: 456].

I gestured in the first chapter at why I think this. Some things — some contents, for example — seem to me to make indispensable contributions to explanations without themselves being *the thing* that explains the *explanandum*. They can, in this way, explain without being *what explains*. Let us all grant that one good explanation of Edmund's skating near the edge of the pond is that he believes that THIN ICE: the ice in the middle of the pond is thin. Let us also suppose that the *explanans* is either Edmund's believing that THIN ICE or the fact that he believes that THIN ICE. In my view, the consideration (or proposition, state of affairs, or content, if you like) that THIN ICE *itself* makes an indispensable explanatory contribution here. One indication of this is that the explanation of Edmund's action, even when routed through Edmund's belief, seems to be hopeless without this content. We could not very well hold that the *explanans* is Edmund's belief that black briefcases are better than brown briefcases, for example. The reason we could not plausibly hold this is that this belief does not have a content suited to figure in an explanation of Edmund's skating near the edge of the pond (at least, not without adding quite a bit more to the story). Once we have observed that the contents of reasons — whatever reasons may be, ontologically speaking — are indispensable to good reasons-explanations of actions, I think we need good reason to refuse the idea that the contents themselves do some of the explanatory work.

I address two natural responses to this way of suggesting that something can explain without being the *explanans*. As we saw in the first chapter, it will do no good to respond here that THIN ICE cannot *by itself* — that is, *in isolation* — do anything to explain Edmund's skating near the edge. For it is not generally true that for some thing Φ to be able to explain some other thing Ψ , Φ must be able to explain Ψ on its own. A better response is that what explains Edmund's skating, for example, is perhaps the compound entity of Edmund's mental state (belief) and the consideration in question (THIN ICE).

When these two things are combined in the right way (namely, when THIN ICE is what Edmund believes), we get something that explains Edmund's action, and which is also the *explanans*. But it is only the whole, and neither of the parts, that explains his action. We can say something similar for the sentences we hold to be *explanantia*. Perhaps "ice" is also indispensable to any good explanation of Edmund's action; but we are not tempted to say that "ice" in any way explains Edmund's action. Rather, it is the whole sentence of which "ice" is a part (thin ice) that explains Edmund's action; and this whole is also the *explanans*.

On this way of thinking I have three things to say. The first is that I am certainly not claiming that every part of every *explanans* itself does some explaining of the *explanandum*. Second, to the extent that we are influenced by this remark about parts and wholes, I suspect it is because we are smuggling in a causal model of explanation. It seems ridiculous to say, for example that the non-obtaining state of affairs THIN ICE *causes* anything at all; but Edmund's belief that THIN ICE surely might cause him to skate near the edge. The problem here is that in giving reasons-explanations of actions, we are not in the business of giving (merely) causal explanations. What we want to know is what motivated, in the rational sense, the agent to do what he did. I am inclined to agree with the Non-Psychologists that, even in cases like Edmund's, what plays this role is THIN ICE rather than Edmund's belief that THIN ICE. Third, even the imposition of a causal model of explanation does not force us to say that, for any whole Φ and part Ψ , when Φ explains some fact p (in being the *explanans* vis-à-vis p), then Ψ does not explain p . It might be the complex of there being sufficient oxygen in the room, the match being dry, and the match being struck on the matchbox that is the *explanans* of the match being lit. But here I find no reason to deny that there being sufficient oxygen in the room at least partly explains (or even causes) the match being lit. Of course, it does not *by itself* or *in isolation* explain the match being lit; but we have dealt with that misguided thought already.

There being sufficient oxygen in the room explains the match being lit only as a part of the whole complex, which is the *explanans*; still, it is true that there being sufficient oxygen explains the match being lit. In an analogous way, I think that we can say that the consideration thin ice explains Edmund's skating near the edge of the pond, even if it only explains as a part of some whole complex which is the *explanans*. Examples of such *explanantia* might include "His reason is that THIN ICE," and "He believes that THIN ICE."⁸⁷

So, I think that we can say, for example, that THIN ICE itself explains Edmund's skating near the edge of the pond. It so explains, and is enabled to so explain, by being the consideration which motivates Edmund to skate near the edge. If we allow that we can give the agent's reason for acting by specifying the consideration which motivated the agent to act, and also that something can explain an action by being the reason for which the action is done, all this really amounts to is a restatement of the axiom of REASONS-EXPLANATIONS. Here we follow Anscombe and Davidson:⁸⁸

The question, "What is meant by the relevant sense of 'Why?'" and "What is meant by 'reason for acting?'" are one and the same.

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

There is a type of action explanation (which I have been calling reasons-explanation) which is simply defined by its explaining via the reason for which the agent acted. Of course, what may be controversial is sticking to this way of thinking when the reason

⁸⁷ There are some who think that, in addition to normative and motivating reasons, there is a separate class of explanatory reasons. Adherents to this taxonomy include Alvarez [2010], Baier [1958] and Hieronymi [2011]; and perhaps also Smith [1994], Darwall [2003], and Mantel [2014]. As Alvarez [2010: 33-39] reveals, though, at least some of the motivation for distinguishing this third class of reasons comes from something like an argument from error I treat here. To the extent that the explanatory framework I offer here is feasible, such motivation is weakened.

⁸⁸ See Anscombe [1957: 9] and Davidson [1963: 685]. See also Setiya [2007: 23, 42].

appears to be a falsehood. Now, if there is independent reason to think that A cannot Φ for the reason that r when it is not the case that r , then that could certainly be brought the bear here. But we examined that controversy in the first section, and found the case for the Factive View_M wanting. Once we have on board that A can Φ for the reason that r even when it is not the case that r , I think that we should stick to our guns and say that in such cases r still plays an explanatory role with respect to A 's Φ ing. This comes down to simply remaining steadfast in our commitment to REASONS-EXPLANATIONS. The subject of controversy should be rather *in what sense* a false motivating consideration can explain an action, and, in the light of our answer to that question, in what sense FACTIVE EXPLANATIONS remains intact. As Bernard Williams observed, there is some pressure not to have our explanatory framework altered for error cases:

If there are reasons for action, it must be that people sometimes act for those reasons, and if they do, their reasons must figure in some correct explanation of their action (it does not follow that they must figure in all correct explanations of their action). The difference between false and true beliefs on the agent's part cannot alter the *form* of the explanation which will be appropriate to his action.⁸⁹

So I think that, if possible, we should reserve the same explanatory role for motivating reasons in error cases as we do in the so-called good cases. What I have been suggesting is that the distinction between playing an explanatory role and being the explanation can help achieve this uniformity.

In general, then, I think it is not true that (i) motivating reasons must always be capable of being the *explanans* of some action; that (ii) in reasons-explanations of actions, the *explanans* must always be identical to the agent's reason for acting; or that (iii) the only way something can explain some other thing is by being the *explanans* of that thing. In general, it is not plausible to suppose that we are only in the business of seeking out an

⁸⁹ See Williams [1981:]. See also Dancy [1995: 426; 2008: 267-268].

explanation if we are in the business of identifying the *explanans*; and it is also not plausible to say that we are only in the business of giving explanations if we are stating truths. Any argument from error in which the reason-robbing move relies on one of these thoughts (or any other thought relevantly similar to EXPLANANTIA ARE REASONS or NO FALSE THAT-CLAUSES) will fail. Such arguments assume or imply implausible constraints on explanation.

ARGUMENTS FROM PROCEDURAL ERROR

So far I have extracted and criticized as implausible at least two possible constraints on action explanation that might underpin arguments from factual error. In this section I want to turn my attention to what I identified earlier as the second group of arguments from error: arguments that focus on cases of procedural error. They focus on cases wherein the agent's shortcoming is *procedural* or perhaps *normative*: even if the agent happens to arrive at an accurate view of the world (and so makes no *factual* error), he has done so in a problematic way. Such is the case for characters like Hornsby's Edmund, who is a familiar sort of character in philosophy. He has arrived at an accurate view of the world — and he even has good reason for holding the view that he does — but intuitively, there is something imperfect about his connection to the way the world actually is. Intuitively, he does not *know* that the ice in the middle of the pond is thin. Now, on some views, this precludes Edmund from acting *for the reason that* the ice in the middle of the pond is thin. Why? Here again is the core of the argument from Hornsby, after having laid out the details of the case sufficient for establishing that Edmund does not know that THIN ICE:

Edmund, then, did not keep to the edge *because* the ice in the middle of the pond was thin... The fact that ice was thin does not explain Edmund's acting, even though

Edmund did believe that it was thin, and even though the fact that it was thin actually was a reason for him to stay at the edge...

As I said earlier, the structure of this argument is not completely clear. But some aspects of how the argument is supposed to work are familiar from the arguments from factual error. It is clear that central to the argument is a judgment about a claim of the form “*P* because *q*.” We cannot truly say, those offering this sort of argument declare, that “Edmund is keeping to the edge *because* the ice in the middle of the pond is thin;” and this is supposed to indicate to us that THIN ICE cannot explain Edmund’s action. But, unlike arguments involving factual error, the ‘because’ claim in question is not false because the *explanans* is false (indeed, it is true). Rather, the ‘because’ claim in question is supposed to be false because the (apparent) *explanans* is not something that Edmund *knows*.

Putting these elements together, here is a rough pass at capturing the line of reasoning here:

[9] In some cases where *A* appears to Φ for the reason that *r*, *A* does not know that *r*.

[10] If *A* does not know that *r*, then *A* cannot Φ because *r*.

[11] If *A* cannot Φ because *r*, then *r* cannot be what explains *A*’s Φ ing.

[12] So, in such cases, *r* cannot be what explains *A*’s Φ ing.

[13] Whenever *A* Φ s for the reason that *r*, *r* can be what explains *A*’s Φ ing.

[14] So, in such cases, *A* cannot Φ for the reason that *r*.

[15] So, in order for *A* to Φ for the reason that *r*, *A* must know that *r*.

Though such arguments are centered around a different kind of error, the basic structure is the same as in arguments from factual error: the first step [9]-[14] is to complete a

reason-robbing inference as to what the agent's reason could not have been, and the second step [14]-[15] is to complete a reason-replacing inference as to what the agent's reason must have been. As I did in the previous section, I would like to focus on the reason-robbing move in this argument. Implicit in this sort of reason-robbing move is something like one of the variations of the Argument from the Factivity of Explanation we critiqued in the first chapter:

[1b] A reason is the kind of thing that can replace q in a claim of the form " P because q ".

[2b] Anything that can replace q in a claim of the form " P because q " is a fact.

[3b] Therefore, all reasons are facts.

To the extent that the reason-robbing argument here does rely on this variation of AFE, it will suffer the criticisms pointed out in the first chapter. Also implicit in the reason-robbing move here is something like the idea that EXPLANANTIA ARE REASONS. This constraint seems to make an appearance in securing claims [13] and [14]. To the extent that the argument here relies on such a constraint, it will suffer the criticisms pointed out in the previous section.

Explanations and Accidents

At this point, though, I should like to focus more on the picture of explanation that seems to be in the background of this Hornsby-style argument. We can begin by re-raising the question as to why anyone should follow the progression from claim [9] to claim [11]. Why should we think that Edmund's not knowing that THIN ICE precludes THIN ICE from explaining his action? After all (in this version of the case), it is true that

thin ice, despite Edmund's not knowing it. So it is not our commitment to FACTIVE EXPLANATIONS that is creating trouble here. Furthermore, even though Edmund does not know that THIN ICE, Edmund is significantly cognitively connected to THIN ICE: he believes it, and perhaps believes it for good reason. Importantly, this is the kind of cognitive connection that we would expect to make a difference to the way that Edmund acts. What intentional actions an agent performs seems to be at least partly determined by what the agent believes. Finally, Edmund is doing exactly what we would expect any rational agent to do, given that he believes that THIN ICE. Indeed, any of us, if asked in place of Edmund why we are skating near the edge of the pond, would likely respond with some appeal to THIN ICE. All of this seems to cast serious doubt on our arrival at claim [11], and on any reason-robbing inference there involved. Intuitively, it does not seem to matter that Edmund does not know that THIN ICE; his merely believing it and being motivated by it seems to put it in a position of explanatory power with respect to his action.

As we saw in the previous chapter, John McDowell gives us a way to see why this Hornsby-style might be appealing. Here again is what I take to be the revealing passage.

... we know that [the agent] takes the relevant fact to obtain, and we know that it does obtain. But in knowing only that much, we do not know whether its obtaining is anything but a happy accident in relation to her cognitive position. [This] version of acting in the light of a fact is not sensitive to the difference between cases in which the obtaining of the fact is, and cases in which it is not, a mere happy accident in relation to the agent's cognitive position... If, but only if, the obtaining of the fact by virtue of whose obtaining the relevant belief is true is *not* a happy accident in relation to the agent's cognitive position, we can say that the fact itself is exerting a rational influence on the agent's will; we can say that in doing what [he] is doing the agent is responding rationally to the fact itself ...⁹⁰

⁹⁰ See McDowell [2013: 16-17].

The suggestion here, it seems to me, is that though a character like Edmund might be cognitively connected to THIN ICE in a significant way, his being connected to *the-fact-that*-THIN ICE seems to be a mere “happy accident.” The accidental nature of this connection compromises the ability of thin ice to play an explanatory role with respect to Edmund’s skating near the edge of the pond — it compromises our ability to explain Edmund’s so skating by casting it as a rational response to THIN ICE.

In the previous chapter I registered doubts about the legitimacy of the notion of *the-fact-itself*. I also began to put pressure on the idea that Edmund’s fortuitous connection to THIN ICE precludes it from being the consideration that motivates him to act as he does. McDowell might be right that Edmund’s being significantly cognitively connected to THIN ICE is in some sense a “happy accident.” But, *given* that Edmund is cognitively connected to THIN ICE in the way that he is — given that he believes it for good reason, and takes it to favor keeping to the edge of the pond — it is certainly no happy accident that he acts as he does. That notwithstanding, McDowell insists that the accidental nature of Edmund’s cognitive connection to THIN ICE prevents it from playing an explanatory role vis-à-vis his action, as something to which Edmund responds rationally. According to McDowell,

There is an intelligible interpretation of the claim that Edmund is keeping to the edge because the ice is thin, on which the claim is not true in [Hornsby’s] story, because [Edmund does not know that the ice is thin], even in the version of the example in which the ice was indeed thin... there is *another* form in which propositions like ‘the ice is (was) thin’ can—sometimes, but not in the case [Hornsby] considers—be brought into explanatory connection with the statements about what people are doing or have done.⁹¹

⁹¹ See McDowell [2013: 18].

This reveals, in combination with the previous passage, that underneath the Hornsby-McDowell-style argument for the Knowledge View is something like the principle that

NO HAPPY ACCIDENTS: in any explanation, the connection between the *explanans* and the *explanandum* cannot be a mere happy accident.

So, for example, in order for a claim of the form “*P* because *q*” or “What explains *p* is that *q*” to be true, the connection between *p* and *q* cannot be a “mere happy accident.” Something like this constraint could explain why the Knowledge View is preferable to the Belief View, since the Belief View allows reasons (and therefore things explanatory of actions done for reasons) to be things *merely* believed by the agent.

Unfortunately, those who might rely on NO HAPPY ACCIDENTS do not do much to explain what it means for a connection to be accidental. There is a way of taking the constraint such that it appears to rule out the possibility of any contingent relation being an explanatory relation. Consider a case of normal causal explanation.

THE VASE AND THE VAN: Vance is in a heated argument with his wife. In a fit of rage, Vance throws a vase — one of his wife's prized family heirlooms — at the wall near where his wife is standing. Vance misses the wall, though, and the vase flies out the window of their third-floor apartment onto the street below. In a bit of bad luck, a van is driving by at just that time; the van runs over the shattered vase, inflicting a flat tire which sends the van out of control into a nearby streetlamp.

It seems clear to me that in this case, it is true that the van crashed into the streetlamp because Vance threw the vase, and that Vance's throwing the vase explains the van's crashing into the streetlamp. It is also clear, though, that it is a mere (un)happy accident that the van crashed into the streetlamp given that Vance threw a vase. After all, Vance was aiming for the wall; the van happened to be driving by just the time that Vance threw the vase; the van was driving just so that one of its tires ran over the shattered vase; and so on. But the accidental nature of the connection between Vance's throwing the vase and the van's crashing into the streetlamp does not give us any pause in saying that the former explains the latter.

So if there is any sense in which NO HAPPY ACCIDENTS is true, it had better not be one which rules out the general possibility that contingent relations might be explanatory relations. I do not think that this is what Hornsby or McDowell intend. Perhaps this kind of argument for the Knowledge View is intended to constrain only the explanation of action in particular. Since we could set up a case just like THE VASE AND THE VAN in which an action replaces the van's crashing into the streetlamp, I think we should also say that the constraint is meant to apply only to the *rational* explanation of action (or reasons-explanations of action). Here another case might be helpful.

THE VASE AND THE VISITOR: Vance is in a heated argument with his wife. In a fit of rage, Vance throws a vase — one of his wife's prized family heirlooms — at the wall near where his wife is standing. Vance misses the wall, though, and the vase flies out the window of their third-floor apartment onto the street below. In a bit of good fortune for a visitor looking for a place to stay, the smashing of the vase draws her attention to a sign indicating that there is a vacancy at the local motel. So, the visitor walks over to the motel.

Though this case includes an action done for a reason, and perhaps even some knowledge on the part of the actor, it seems to me relevantly similar to THE VASE AND THE VAN in at least the following ways. We can explain the visitor's walking over to the motel by saying that she walked there because Vance threw a vase. But of course, the connection between these two is utterly accidental. We can also explain the visitor's walking over to the motel by citing her reason, which is perhaps that the "vacancy" sign was illuminated (or even just that *there was a vacancy*). It is true that the visitor walks over *because* the sign is illuminated, even though her knowing that the sign is illuminated is a (happy) accident. In the same vein, we can say that the visitor walks over because there is a vacancy at the motel.

So we have found another understanding of no happy accidents such that it is not plausible: the fact it is an accident that the agent knows that r does not preclude r from being able to explain the agent's action, even as the agent's reason for acting. Still, though, I do not think we have arrived at the sense in which Hornsby and McDowell might endorse no happy accidents. Here is a case which might point more toward their sense.

THE VASE AND THE VACANCY: Vance is in a heated argument with his wife. In a fit of rage, Vance throws a vase — one of his wife's prized family heirlooms — at the wall near where his wife is standing. Vance misses the wall, though, and the vase flies out the window of their third-floor apartment onto the street below. In a bit of good fortune for a visitor looking for a place to stay, the smashing of the vase draws her attention to a sign indicating that there is a vacancy at the local motel. So, the visitor walks over to the motel, and upon inquiring, acquires one of the

vacant rooms. Unbeknownst to the visitor, though, the “vacancy” sign is only illuminated because the motel owner accidentally flipped the switch when leaning against the wall.

We saw in THE VASE AND THE VAN that we could say that Vance’s throwing the vase explains the van’s crashing even though the connection between the two is highly contingent. We then saw in THE VASE AND THE VISITOR that we could say that Vance’s throwing the vase explains the visitor’s walking over to the motel, even though that connection was also highly contingent. Furthermore, we could also say that the visitor walks over to the motel because the sign is illuminated, or even because there is a vacancy at the motel. But can we say the same in THE VASE AND THE VACANCY? There is certainly some awkwardness in saying that the visitor walks over to the motel *because* there is a vacancy; and so we might also think that we should not say that there being a vacancy *explains* the visitor’s walking over. The best account of this awkwardness might be, as Hornsby and McDowell diagnose, the visitor’s fortuitous connection to *the fact* that there is a vacancy — her cognitive connection to that fact does not amount to knowledge. It is a mere happy accident that she *truly* believes that there is a vacancy.

However, to the extent that these considerations make us hesitant to say that the visitor walks over *because* there is a vacancy, I suspect that this is because we are once again smuggling in a causal conception of explanation. Since the vacancy *itself* is no part of the causal chain bringing about the visitor’s walking over, it does indeed seem incorrect to say that the fact that there is a vacancy *causes* the visitor to walk over to the motel. But as we have observed several times now, in wondering whether the consideration that there is a vacancy could, *qua motivating reason*, explain the visitor’s walking over, we were never wondering what *caused* anything. We were wondering what consideration motivated

the visitor to do what she did — what consideration cast walking over in a favorable light for the visitor. Indeed, as Hornsby admits, the ‘because’ of interest “can be glossed with ‘for the reason that’;” and we are to understand such claims to “rule out cases where the ‘because’, of ‘because *p*’ and ‘because they believed that *p*’, is not the ‘because’ of a reason-explanation.”⁹²

So we should sift out the claim that “The visitor’s walking over was caused by there being a vacancy” from the claim that “The visitor walked over because there was a vacancy,” and rather focus only on the sense of the ‘because’ claim that could be rendered as “The visitor walked over for the reason that there was a vacancy.” Once we have done this, though, it seems completely correct to say that the visitor walked over *because* there was a vacancy; and it is not at all clear what difference lack of knowledge on the part of the visitor should make. Those making the Hornsby-McDowell-style argument for the Knowledge View might like to say something like:

If the visitor knows there is a vacancy, we can say that it is no accident, given that there is a vacancy, that the visitor walks over to the motel.

Though this is true, it gives us no grounds on which to perform a reason-robbing move on the visitor who merely believes that there is a vacancy. For what we are interested in is the explanatory connection between *there being a vacancy* and *the visitor’s walking over*. I take it no one would want to claim, in such a case, that the explanatory relation between these two things holds all on its own. Rather, the explanatory relation is enabled by a cognitive fact about the visitor (namely, that the visitor knows that there is a vacancy). The sense in

⁹² See Hornsby [2008: 251-252].

which the explanatory relation is enabled is not that it makes it possible that there being a vacancy causes the visitor to walk over, but rather that it makes it possible that there being a vacancy is something which could have served as the visitor's reason for walking over. But with this in view, it seems clear that defenders of the Belief View could, in just the same way, say

If the visitor believes there is a vacancy, we can say that it is no accident, given that there is a vacancy, that the visitor walks over to the motel.

As above, one making this claim need not hold that there is some bare, unsupported explanatory relation between there being a vacancy and the visitor walking over to the motel. Rather, the explanatory relation is enabled by a cognitive fact about the visitor (namely, that the visitor believes that there is a vacancy). Just as above, the sense in which this explanatory relation is enabled is the *rational* sense. It will of course do no good to object at this point that in cases like THE VASE AND THE VACANCY, it is *the visitor's believing* that there is a vacancy (and not there being a vacancy) that really explains her walking over — since just the same observation could be levied against the account involving knowledge.

So it seems clear that there is a legitimate sense in which things merely believed by agents can, when they motivate agents to act, stand in an explanatory relation to the actions of those agents. McDowell is one who seems happy to concede this. But he also wants to reserve a special (that is: distinct, and perhaps elevated) role for the case in which the agent *acts in the light of a fact*. As we have seen, it is not clear what grounds there is for distinguishing acting in the light of a fact (where “fact” refers to “the-fact-itself”) from acting in the light of a consideration which is a fact. Before putting aside NO HAPPY

ACCIDENTS as implausible, though, I want to consider one more way that McDowell tries to establish the distinction. I think it is the examination of this final point that will elucidate why no argument from error based on a thought like NO HAPPY ACCIDENTS is likely to give advantage to the Knowledge View over the Belief View. McDowell says that

Even though it is a fact, in that version of the example, that the ice in the middle of the pond is thin, we cannot explain his keeping to the edge by saying that that fact is weighing with him... or that in keeping to the edge he is responding to that fact... But suppose we change the example to one in which... Edmund can count as knowing... Now we can explain his keeping to the edge as his rational response to the fact that the ice is thin.

There is no need to deny that an explanation in one of the forms [the Belief View] countenances can provide some understanding of an action. But if we have only that understanding, we do not yet know the answer to a question that should concern us if we are interested in how the action manifests the agent's practical rationality at work. We do not yet know, and we ought to want to know, whether the action can be understood as a rational response to the fact in question.

We do not have to choose between the non-factive and knowledge-requiring forms, as if each were a candidate for capturing the whole truth about how we understand actions when we see them as the agent's rationality at work. The point is that if... we know only something we can express in one of the forms [the Belief View] explains, we know some, by all means, but only some of what there is to know about the rational intelligibility of the action we are considering. There is a question about its rational intelligibility that we are not yet in a position to answer: the question whether the agent, in acting as she does, is rationally responding to the fact in virtue of which her relevant belief is true.⁹³

There is much to like and take on board in McDowell's insightful remarks here. Summarizing, I take McDowell to be making at least the following points. First, there is no flaw, *per se*, in the reasons-explanations that explain *A*'s Φ ing by appeal to a motivating consideration which the agent merely believed. Second, these reasons-explanations are compatible with those that require the agent to know her reason. Third, there is a

⁹³ See McDowell [2013: 19-20].

significant sense in which if A does not know that r , we cannot explain A 's Φ ing by casting it as a practically rational response to r .

I make no complaint about the first two points. Instead, I take issue with the third point. As I said in chapter one, it seems clear that we can indeed say that, in skating near the edge of the pond, Edmund is responding rationally to the consideration that THIN ICE. This seems to me undeniable; and whether Edmund knows that THIN ICE, or indeed whether it is even true that THIN ICE, is irrelevant to this fact. Insofar as casting something as a rational response implies an explanatory relation, as McDowell seems to hold, then it is also clear that we can place THIN ICE in an explanatory relation to Edmund's skating. McDowell claims, though, that if Edmund does not *know* that THIN ICE, then the fortuitousness of the connection between THIN ICE and Edmund's action somehow prevents us from understanding Edmund's action as "manifesting his practical rationality at work." Furthermore, this fortuitousness also prevents us from knowing everything there is to know about the "rational intelligibility" of Edmund's action. So, while there may be some explanatory role to be played by motivating considerations merely believed to be the case, there is a further explanatory role only to be played by motivating considerations known to be the case.

At this point, I think we need to acknowledge that we can (and should) separate the rational intelligibility, explanation, and even evaluation of Edmund's skating near the edge of the pond from that of his believing that the ice in the middle of the pond is thin. This separation has nothing to do with the oft-supposed divide between belief and action, or between so-called "practical" and "epistemic" rationality. The separation is rather the result of simply recognizing that *skating near the edge* and *believing that the ice is thin* are two different things done by Edmund, as responses to different considerations. Our account of the latter need not infect our account of the former; and the rational intelligibility of the

latter makes no difference to that of the former. We need not know anything about why or how Edmund came to take that THIN ICE. Once we know that he does so take it, we are in a position to explain his action by casting it as a response to THIN ICE. We can of course say that, in some sense, the case in which A Φ s for the reason that r , A knows that r , and A 's Φ ing is the proper response to the consideration that r , is the exalted case of acting for a reason to which we all aspire. But it is difficult for me to locate any sense in which adding that A knows that r to the fact that A Φ s for the reason that r somehow makes A 's Φ ing *more* of a case of acting for a reason. McDowell claims that we can see that certain accounts of the rational intelligibility of Edmund's action are missing something by the fact that, if Edmund does not know that THIN ICE, we cannot on those accounts answer the question as to whether Edmund is rationally responding to the fact in virtue of which his belief that THIN ICE is true. But what distinguishes this question from the question of whether Edmund rationally responds to the consideration that THIN ICE, which is a fact that he believes? One might want to say that there is a difference between responding to the consideration that THIN ICE and responding to *the truth of the consideration* that THIN ICE (or *its being the case that* THIN ICE). But insofar as there is a significant difference here, on all accounts Edmund responds to the former and not the latter. Anyway, I suspect that this distinction is not quite what a McDowell-style Knowledge View has in mind. Ultimately, I instead suspect that the distinction between my two questions, according to this kind of Knowledge View, is propped up by the supposed distinctness of the mysterious notion of *the-fact-itself*.

Summary Remarks

We began this section conjecturing that something like NO HAPPY ACCIDENTS might underpin certain reason-robbing moves in arguments from error, which are

supposed to teach us that there is an important sense in which A cannot Φ for the reason that r unless A knows that r . On some understandings, NO HAPPY ACCIDENTS is implausible just as a constraint on explanation. On at least one of the plausible renderings, its plausibility relies on restricting ourselves to a purely causal notion of explanation. Once we notice these two facts, we struggle to find an understanding of NO HAPPY ACCIDENTS that could aid us in completing the reason-robbing move supposed to indict the Belief View and pressure us toward the Knowledge View. While it seems true that there is a difference between the cases where the agent knows her reason and merely believes her reason, the difference seems wholly to do with the agent's cognitive state rather than her acting.

VARIATIONS IN EXPLANATION

At this point we have examined at least three constraints on explanation which certain arguments from error might rely on. In this section I want to consider two more constraints on action explanation that tend to hover even farther in the background. While principles like NO FALSE THAT-CLAUSES, REASONS ARE EXPLANANTIA, and NO HAPPY ACCIDENTS seem useful only to some particular arguments from error, I suspect the two constraints examined in this section are much more widely relied on.

Translatability

I start with that of the two constraints more relevant to the material of the previous section. Recall that one argument from error goes as follows.

[1b] A reason is the kind of thing that can replace q in a claim of the form “ P because q ”.

[2b] Anything that can replace q in a claim of the form “ P because q ” is a fact.

[3b] Therefore, all reasons are facts.

We saw that something like this argument might be thought to favor the Factive View_M, Psychologism, or even the Knowledge View. In the first chapter, I resisted claim [1b] on the grounds that it implies that playing an explanatory role can be reduced to playing a role in a ‘because’ claim. In doing so, I anticipated a kind of resistance based on the following remarks from John Hyman and Constantine Sandis.⁹⁴

Reasons can be stated or given; and the canonical form of a sentence stating or giving a person’s reason for doing or believing something is ‘*A* Φed because *p*’.

All reasons-statements must be translatable (without change of meaning) to any of the standard forms of explanation, on pain of failing to qualify as explanatory.

In conjunction with the argument above, these considerations might be thought to exert some real pressure toward Factive View_M, Psychologism, or the Knowledge View. If, whenever *A* Φs for the reason that *r*, there must be a true explanatory claim of the form “*A* Φs because *r*,” the axioms of REASONS-EXPLANATIONS and FACTIVE EXPLANATIONS threaten to all but rule out views like the Nonfactive View_M, Non-Psychologism, or the Belief View (by some kind of argument from error).

There are two issues raised by remarks such as Hyman’s and Sandis’. The first concerns what is the *canonical* form of reasons-explanation; the second concerns what we might call the principle of

TRANSLATABILITY: for any two legitimate forms of reason-explanation, it must be possible to translate between them without any change of meaning.

⁹⁴ See Hyman [1999: 443] and Sandis [2013: 38], respectively.

To get a fuller idea of what this constraint on action explanation entails, here is some more of the context for the Sandis passage above.

There is an obvious objection to my line of argument so far, namely that when asked to explain our actions we tend to simply state the considerations we (take ourselves to have) acted upon. Similarly, when we explain the acts of others we frequently do so by directly stating their reasons, as follows:

(b) He did it because it would increase his pension.

Are these not instances of explaining an action by stating the reason itself and not through some further statement about what one's reasons are? Not if (b) is elliptical... as John Hyman supposes... Dancy (2011) acknowledges the 'factive pressure of the word "because",' but denies that this factive way of giving a reasons-explanation exhausts the possibilities, rightly complaining that 'the election of one rather than another way of giving a reasons-explanation as canonical seems [to be] arbitrary and tendentious'. But the fact that no one form is canonical does not entitle us simply to ignore the form given above, for it is not a question of *favoring* the factive form over the teleological (or vice-versa). Rather, all reasons-statements must be translatable (without change of meaning) to any of the standard forms of explanation, on pain of failing to qualify as explanatory.

Now, I am inclined to agree with Dancy that debates over the canonical form of reasons-explanations are not likely to be fruitful or enlightening. But if Sandis (or the principle of TRANSLATABILITY) is correct, which form is canonical might make no difference. Indeed, views about which explanatory forms are canonical might be completely irrelevant without some such constraint. On this way of thinking, once something is established as at least being a *standard* form of reasons-explanation, any additional candidate form can be (partly) tested for legitimacy by whether it can be translated into that standard form.

One can see how this principle might be used to eliminate explanatory resources useful to opponents of the Factive View_M, Psychologism, or the Knowledge View. In addressing the argument given by [1b]-[3b], for example, I suggested in the first chapter that one might refuse [1b] in favor of

[1b*] A motivating reason is the kind of thing the content of which can replace r in a claim of the form, “ A ’s reason for Φ ing is r .”

The truth of [1b*] might be thought to do damage to, say, the Factive View_M , since if [1b*] is true, one cannot plausibly use FACTIVE EXPLANATIONS to argue that reasons are factive. For the Nonfactive View_M on offer might hold that the explanation of A ’s Φ ing is the (true) claim that A Φ s for the reason that r . Though there may here be a violation of a principle like NO FALSE THAT-CLAUSES, there is no violation of any plausible understanding of FACTIVE EXPLANATIONS. But is the claim that “ A Φ s for the reason that r ” a legitimate reasons-explanation of A ’s Φ ing? Perhaps not, if the principle of TRANSLATABILITY is correct. For we know that “ A Φ s because r ” is a standard explanatory form; but we may not be able to translate a claim of the form “ A Φ s for the reason that r ” into such a ‘because’ claim without any change in meaning. A hint that this is so is that, on their presently intended uses, one of these forms requires that r be true, while the other does not. But if we cannot translate “ A Φ s for the reason that r ” into the standard “ A Φ s because r ,” then by the TRANSLATABILITY constraint the former claim cannot be counted among the legitimate explanatory forms. Furthermore, since we only thought of “ A Φ s for the reason that r ” as an explanatory claim due to the thought that we can explain A ’s Φ ing by giving the reason for which A Φ d, perhaps we should walk this line of thought all the way back and reject the idea that it can even be true that A Φ s for the reason that r when it is not the case that r .

I think that this way of reinforcing various arguments from error is flawed. I give two reasons why. The first is built on considerations I raised in the first chapter. It seems to me undeniable that at least one standard reason-*giving* form is: A Φ s for the reason that r . As I said, it might simply be an analytic truth that claims of this form give agents’

reasons for acting. Combining this fact with the axiom of REASONS-EXPLANATIONS, it seems hard to deny that such reason-giving claims are also in the business of giving reasons-explanations. Denying this would seem to commit us to denying that we can explain an action by giving the reason for which it is done. For if “ A Φ s for the reason that r ” is a standard reason-giving form, and we can explain an action by giving the reason for which it is done, it seems to follow that “ A Φ s for the reason that r ” should also be considered a standard form of reasons-explanation. I suspect that an analogous line of reasoning will be available for many other explanatory forms not friendly to arguments from error, such as

An answer to the question of why A is Φ ing is: for the reason that r .

The consideration that r is what motivated A 's Φ ing.

The reason for which A is Φ ing is that r .

In Φ ing, A is responding to the consideration that r .

A is Φ ing in light of the consideration that r .

I think that these should all be considered standard forms of reasons-explanations — and unobjectionably so, in part due to reasoning of the sort just laid out. They are all standard forms of reason-giving claims. If these claims are not inter-translatable either among themselves or with other standard reason-giving forms (such as A Φ s because p), then we should say: so much the worse for TRANSLATABILITY.

The second flaw in propping up an argument from error with thoughts about translatability is that the principle of translatability seems to me implausible even independent of considerations about standard reason-giving forms. To see this, we should first observe that, on any plausible view, the list of *standard* forms of reasons-explanation is

surely longer than one entry. I take it that no one wants to say that the only legitimate form of reasons-explanation is “ P because q .” Instead, any complete and agreeable list is likely to include at least:

A Φ d because p .

What explains A 's Φ ing is that p .

The reason why A is Φ ing is that p .

More controversially, we might add:

P is what caused A 's Φ ing.

A Φ d because A believed that p .

A Φ d because A wanted to Ψ .

Now, there is already a serious question about whether TRANSLATABILITY rules out some of these forms (and if so, which ones can be ruled out non-arbitrarily). But for a clear-cut example, we can attend to the earlier thoughts from Scott Sehon and G. F. Schueler (respectively), and also from Arthur Collins:⁹⁵

[Common-sense] explanations [of action] are not causal; instead, I claim that they are *teleological*. A teleological explanation explains by citing the *purpose* or *goal* of the behavior in question... The paradigmatic form of a teleological explanation is A Φ d in order to Ψ .

Explanations of actions in terms of the agent's reasons, I want to claim, work by citing the purpose or purposes for which the person who performed the action actually acted.

⁹⁵ See Sehon [2005: 13] for what comes before the ellipses and [2005: 149] for what comes after; see Schueler [2003: 56] for the second passage, and Arthur Collins [1984: 315-316] for the third.

Those explanations that Davidson called “rationalizations,” the explanations for which he said that “primary reasons” can be given, appear to fit the pattern of teleological [i.e. purposive] explanation quite readily.

These remarks suggest — and I am inclined to agree — that one standard form of a reasons-explanation is: $A \Phi$ d in order to Ψ .

Now, can we reliably translate claims of the form “ $A \Phi$ d because p ” into claims of the form “ $A \Phi$ d in order to Ψ ” without any change of meaning? I think not. It seems obvious that, whether or not we can translate infinitival clauses into that-clauses (as we considered in the first chapter), the following two claims simply do not have the same meaning.

Edmund skates near the edge because he believes there is thin ice in the middle.

Edmund skates near the edge in order to avoid thin ice in the middle.

In any attempt to translate, the fact that both claims include the clause “Edmund skates near the edge” seems a promising start. But despite the great deal these two claims have in common, I cannot see any principle of translation, which does not amount to the bald assertion of a brute rule, that would allow us to translate between them. By what rule might we translate “because” into “in order to” or “he believes there is thin ice in the middle” into “avoid thin ice in the middle”? Even if we thought that, for example, that the “in order to” clause was somehow elliptical for something like, “his purpose is to avoid thin ice in the middle,” this would not help. For it is no more plausible that

“Edmund skates near the edge because he believes there is thin ice in the middle”

is a legitimate translation of

“Edmund skates near the edge because his purpose is to avoid thin ice in the middle,”

or vice versa. Given the great deal of commonality, we might have thought this should be an easy case of translation — much easier, anyway, than examples like

Edmund skates near the edge because he believes there is thin ice in the middle.

Edmund skates near the edge in order to avoid drowning.

It seems to me clear that there is little to no hope of performing any translation between these two explanatory claims. But that does not in any way lessen our confidence that they are both perfectly adequate *qua* explanatory claims.

One might object that, as in some other situations, my critical example has been purposive, and that purposes are not reasons. I myself find it somewhat appealing to hold that reasons and purposes belong to the same ontological category. But the ontology is irrelevant here. The implication of this objection is that purposive explanations are in the business of identifying purposes; and since purposes are not reasons, then purposive explanations should not be counted as reasons-explanations (much less as *standard* reasons-explanations). However, as I argued in the first chapter, reasons-explanations are not necessarily in the business of identifying something to be *equated* with the agent’s reason for acting. It is no more plausible that purposive explanations are necessarily in the business of identifying something to be equated with the agent’s purpose (and therefore, *not* the agent’s reason). Furthermore, as I also pointed out in the first chapter, the (purported) fact that purposes are not reasons does not entail that claims of the form “A

Φ d in order to Ψ ” are not reason-giving. Indeed, it seems undeniable that “ A ’s reason was to Ψ ” is a standard reason-giving claim. So, by the same reasoning as given earlier, it seems to follow that such claims should also be counted as reasons-explanations. Anyway, it is independently plain to see that purposive explanations are in the business of doing exactly what reasons-explanations are: they reveal the good the agent saw in doing what she did, and specify the considerations or features that motivated her to act as she did.⁹⁶ With these considerations in mind, I cannot see why any claim about an ontological difference between reasons and purposes should lead us to believe that purposive explanations are not reasons-explanations.

Perhaps there is some suspicion that the difficulty here is created by the fact that purposive explanations include infinitival clauses, while the other standard forms do not. In view of that, one might want to endorse a more modest TRANSLATABILITY claim; for example, one might want to claim only that we should be able to translate between any two standard *propositional* reasons-explanations. This might help defenders of the Factive View_M, Psychologism, or Knowledge View sidestep any opposing considerations based on purposes, purposive explanations, or infinitival clauses. A Factivist might concede, for example, that there can be no factive constraint on infinitival clauses. After all, it is not even clear that infinitival clauses accept truth-values. But, so long as we are in the business of giving reasons or reasons-explanations in propositional form, the propositions picking out the agents’ reasons must be true. The more modest TRANSLATABILITY constraint might be applied here, and pressure us away from the possibly non-factive “ A ’s reason is that r ” by appeal to the certainly factive (and standard) “ A Φ d because p .”

⁹⁶ See Collins [1984: 324] for a similar thought, although I do not agree with Collins that “reason-giving explanations obviously refer actions to objectives, outcomes, or end-states that are caused by the explained events.

Even this more modest TRANSLATABILITY is implausible, though. For we need not use an infinitival clause to pick out a reason that is purposive, goal-directed, or instrumental in nature. — and that seems to me more the cause of the difficulty for TRANSLATABILITY than any particular packaging of a reason or reasons-explanation. Consider, for example:

Edmund skates near the edge because he believes there is thin ice in the middle.

Edmund skates near the edge so that he will not fall through thin ice in the middle.

I think that, as with purposive accounts, the second statement here is clearly reason-giving and explanatory. The second statement is also propositional: the ‘so that’ specifies a motivating consideration in propositional form, just as the ‘because’ operator does. But it seems to me no easier to translate between this pair of claims than it was to translate between the previous pair. We might try, for example:

He is doing that because he believes that if he does, he will not fall through the ice.

He is doing that so that he will not fall through thin ice in the middle.

These two claims look more similar, such that we may have higher hopes of translating between them. But do they have the same meaning? Not necessarily. For in giving a reasons-explanation, we hope to reveal the consideration that cast the action in a favorable light for the agent. Edmund surely could have thought to himself, “I had better stay near the edge so that I do not fall through” without thinking “If I skate near the edge,

I will not fall through.” Edmund might merely have considered the risk without considering what it would take to actually suffer the dreaded consequences; he might be aware that local bullies may come toss him in even if he keeps to the edge; he may think only that skating near the edge lessens his chances of falling through; and so on. So it seems to me that the above pair of claims do not even *imply* one another — much less that there is a way of translating between them without change of meaning.

The foregoing lead me to think that efforts to establish either a definitive and exclusive list of canonical or standard forms of reasons-explanations, or an explanatory constraint like TRANSLATABILITY, are misguided. There are simply too many regular, familiar, and seemingly acceptable reason-giving forms for such tactics to account for; and there are some particular forms that pretty clearly throw a wrench in things. I think it difficult, if not impossible, to rule out on such grounds any of the following forms as genuine reasons-explanations.

$A \Phi$ s for the reason that r .

A 's reason for Φ ing is that r .

A is responding to the consideration that r .

$A \Phi$ d because p .

A is Φ ing because A believes that p .

A is Φ ing because A 's reason is that r .

$A \Phi$ d because A wanted to Ψ .

$A \Phi$ d in order to Ψ .

A 's purpose in Φ ing is to Ψ .

A is Φ ing in view of Ψ .

A is Φ ing so that p .

I make no claim about whether these reasons-explanations amount to the same explanation in every particular instance. I also make no claim about the ontologies of the entities ostensibly referenced by these explanatory claims, except to say that I think each of these claims may plausibly be understood as specifying A 's reason for Φ ing. So long as we accept an understanding of the axiom of REASONS-EXPLANATIONS such that we can explain an action by specifying the agent's reason for performing that action, I think any of these forms might count as legitimate reasons-explanations of an action in a particular case. Any argument from error that relies on ruling out any of these explanatory forms with the sort of reasoning critiqued in this section, then, is unlikely to succeed.

Propositionalism

The final explanatory constraint I would like to examine is one which I am inclined to think that everyone on all sides of the arguments from error assumes to be true. As such, my rejection of this explanatory constraint would be perhaps the most radical of those I consider. I will call this constraint

PROPOSITIONALISM ABOUT REASONS-EXPLANATION: whenever an agent acts for a reason, any reasons-explanation of that action must be a proposition.

I assume that it is transparent how an assumption like this has been part of the backdrop of the discussion so far. It is perhaps a result of combining (a certain interpretation of) REASONS-EXPLANATIONS and the natural thought that reasons are propositional entities. But it has independent plausibility of its own. After all, we have seen that many of the standard reason-giving and reasons-explanation forms are propositional, at least in that a

that-clause is often used to pick out the motivating reason or the *explanans* of the action. Indeed, this observation is not limited to reasons-explanations: the very general explanatory forms “*P* because *q*” and “What explains *p* is that *q*” are both propositional.

By the same token, it is also not difficult to see how the implausibility of PROPOSITIONALISM would create some complications for many views and arguments in the area. On the additional assumptions that anything capturable by a that-clause is a proposition, and that the contents of mental states are propositional in this sense, it would become at least awkward to see how the Factive View_M, Psychologism, the Knowledge View, and even the Belief View should be articulated (much less how they could be true). Similarly: on the assumption that propositions are the sort of thing that can be true or false, if reasons-explanations need not be propositions, it likewise becomes more difficult to see how any argument from error could rob Edmund, for example, of his reason for skating near the edge on the grounds that only truths can explain things. This is only one example; I trust it is clear how other arguments in the area might be similarly complicated.

There is, then, a significant question about whether and in what sense explanation is propositional. This question is of interest to frameworks for explaining action, but also to explanatory frameworks in general. I shall not here defend the general or unqualified claim that explanation is not propositional. Instead, I wish to attend to some considerations which I think at least put some pressure on PROPOSITIONALISM as an explanatory constraint, and perhaps encourage us to articulate it more precisely. A starting point here is to wonder to what extent it is generally plausible that explanation is propositional. A starting point *there* is to wonder about the ontology of *explanantia*. As we saw in the first chapter, it is perhaps natural to think of explanations as propositions or sentences, à la Hempel and Oppenheim:

By the explanandum, we understand the sentence describing the phenomenon to be explained (not that phenomenon itself); by the explanans, the class of those sentences which are adduced to account for the phenomenon.⁹⁷

This is natural enough. But there are other ways of conceiving of explanatory entities. We might think, for example, that states of affairs can explain things — even actions.

According to Jonathan Dancy,

Like Plantinga, I see an ontological gulf between things capable of being the case (i.e. states of affairs) and things capable of being true (either propositions or sentences). And only those capable of being the case are capable of being a good reason.⁹⁸

If we combine this thought with Dancy's thought that "it must be possible to act for a good reason," as well as our axiom of REASONS-EXPLANATIONS, it might seem sensible to suppose that it is not propositions but rather states of affairs that explain actions done for reasons.

Another common thought is that some of the things we wish to explain are events, and events are explained by other events. But events do not seem to be propositions. Even if, as Roderick Chisholm puts it, there is an intimate connection between propositions and states of affairs, we might think propositions and events are more distantly related:

Events and propositions are species of states of affairs. A *proposition* could be defined as any state of affairs which is necessarily such that either it or its negation does not occur... We could now say that an *event* is a contingent state of affairs which is not a proposition and which implies change.⁹⁹

Indeed, we might think, with Davidson, that the entities that explain actions in particular are not just any events, but those events which caused the action:¹⁰⁰

Events The most primitive explanation of an event gives its cause.

⁹⁷ See Hempel and Oppenheim [1948: 136-137].

⁹⁸ See Dancy [2000: 117] for the passage and [2000: 103] for the quote connecting normative and motivating reasons.

⁹⁹ See Chisholm [1970: 20]. For more on the events-based conception of action explanation, see Davidson [1980], Hornsby [2004], and Kent Bach [1980].

¹⁰⁰ See Davidson [1963: 698, 695].

The signaling driver can answer the question ‘Why did you raise your arm when you did?’, and from the answer we learn the event that caused the action.

Or, as Anscombe observed, we might think that even the view that actions are explained by those events which cause them admits of further disambiguation:

If the question ‘Why did you do that?’ is put by someone who makes it clear that he wants me to mention the mental causes—*i.e.*, what went on in my mind and issued in the action—I should perhaps give this account.¹⁰¹

We could go on here. But there is a *prima facie* case to be made that, ontologically, explanatory entities need not be propositions. So, if by PROPOSITIONALISM we intend to declare that the entity that explains an action must, ontologically speaking, be a proposition, this claim faces a steep intuitive burden.

However, as I have argued, our explanatory frameworks need not be limited by the ontology of explanatory entities. Even if what ultimately explains (in some metaphysical or causal sense) an action is an event, for example, we may be able to characterize that event propositionally (for example, “Angus punched his boss because his boss fired him”). In that case, the explanation of the action may still be propositional in a significant sense. Furthermore, in examining the merits of PROPOSITIONALISM, we are not simply interested in the nature of explanation in general (though that is of interest); rather, we are trying to focus in on constraints particular to action explanation. So perhaps these remarks about the possibly varying ontologies of explanatory entities in general are not directly to the point.

It may be that there is something unique about reasons-explanations of actions, such that their form should be propositional. As I mentioned before, this thought could result from combining REASONS-EXPLANATIONS with the natural thought that reasons are

¹⁰¹ See Anscombe [1957b: 324].

propositional entities. After all, of the forms of reason-giving claims I put forth in the previous section, many of them do explicitly capture reasons in that-clauses:

A Φ s for the reason that *r*.

A's reason for Φ ing is that *r*.

A is responding to the consideration that *r*.

A is Φ ing because *A* believes that *p*.

A is Φ ing so that *p*.

Using the thought (which I have advocated for) that we can explain an action by offering up a reason-giving claim, this might lead us to believe that there is some significant or general sense in which PROPOSITIONALISM is true. But is it true that reasons-explanations *must* proceed in this way? This is not so clear to me. A starting point for putting pressure on this thought can be provided by some of the points I have already made in treating explanatory claims which I have called purposive and fragmentary. In addition to those just mentioned, the following reason-giving forms also seem adequate:

for the reason that *r*.

because *p*.

because *A* believes that *p*.

in order to Ψ .

in view of Ψ .

so that *p*.

Each of these, I think, would be perfectly acceptable ways of answering a question of the form, “Why is A Φ ing?” This is in no small part due to the fact that they each seem to be acceptable ways of making known that which motivated A to Φ , or which cast Φ ing in a favorable light for A . This would seem to suggest that these clauses are acceptable reasons-explanations, and therefore that PROPOSITIONALISM is in some significant sense false. It does not seem true that explanatory utterances, clauses, or claims must be offered in propositional form.

One could object here that at least some of these fragmentary accounts do not offer up *the fragmentary clause* as the *explanans*. When we say, for example, that “She did it because her boss fired her,” we offer up not “because her boss fired her,” but rather “her boss fired her” as the *explanans* of her action. But “her boss fired her” is a proposition; so the “because p ” answer to the question “Why?” does not cast doubt on PROPOSITIONALISM in any serious way. Furthermore, of those fragmentary accounts for which we cannot say this, we can say that each of these fragments anyway seem to be elliptical for complete sentences or propositions:

A is Φ ing for the reason that r .

A is Φ ing because p .

A is Φ ing because A believes that p .

A is Φ ing in order to Ψ .

A is Φ ing in view of Ψ .

A is Φ ing so that p .

If anything, we should say that *these* are the explanations of A 's Φ ing (and so, again, there is no serious challenge to PROPOSITIONALISM here).

As in the first chapter, we need not address the plausibility of the thought that these fragments are really elliptical for propositions or sentences. There are at least two points of response that do not require wading into that issue. The first is that even if these fragments are somehow more properly conceived of as elliptical propositions, it would not follow that what these fragments offer as explaining the action is a proposition. That is: it would not follow that what is playing an explanatory role is propositional. In fact, if we consider the line of objection here, we find more evidence supporting the idea that what is being offered to explain the action is *not* a proposition. For nothing is added to these fragments, *qua* explanatory claims, by transforming them into propositions. These propositions do not better enable us to locate *A*'s reason for Φ ing, understand the consideration that motivated *A* to Φ , or lend *A*'s Φ ing a greater level of rational intelligibility. On these fronts, we are just as satisfied with the fragmentary answers to the question "Why is *A* Φ ing?" Furthermore, if a reasons-explanation can be something that explains an agent's action by giving the agent's reason for acting, then (as we have been saying) it seems any reason-giving claim should count as a reasons-explanation. The fragmentary accounts do just as well as the propositional accounts *qua* reason-giving claims. If we ask Edmund why he is skating near the edge of the pond, and he responds, "For the reason that the ice in the middle of the pond is thin," it is not as though we are at a loss as to what his reason for acting is. Even if this answer did somehow leave us at such a loss, we would not be helped at all if Edmund tried to clarify by saying, "My reason for skating near the edge is that the ice in the middle of the pond is thin." These things being so, we should resist the demand that these fragments be transformed into the corresponding propositions before we count them as legitimate explanations. It seems that nothing by way of explanation would be gained in that transformation.

A second point of response was raised earlier in this chapter, and depends on the distinction I tried to draw between something's *explaining* and something's *being the explanation* or *explanans* of some other thing. Let us suppose that "to Ψ " or "for the reason that r " cannot be *the explanation* or the *explanans* of A 's Φ ing. From this it does not follow that these things cannot in any way explain A 's Φ ing. Even if the clause "for the reason that there is thin ice in the middle of the pond" can only rationally explain Edmund's skating as a proper part of the larger complex, "He did it for the reason that there is thin ice in the middle of the pond," it still would not follow that the fragmentary clause does not explain Edmund's skating. We can always say that these clauses explain Edmund's action by giving or specifying that which motivated Edmund to so act.

It may finally be objected that I have relied heavily on the conception of reasons-explanation given earlier by such characterizations as

[1a] A reason is the kind of thing that can be given in an answer to a "Why?" question.

[1d] A reason is the kind of thing that can be a part of an action explanation.

[1e] A reason is the kind of thing that can contribute to the explanation of an action.

and have implicitly rejected those given by such characterizations as

[1b] A reason is the kind of thing that can replace q in a claim of the form " P because q ".

[1c] A reason is the kind of thing that can be the *explanans* of an action explanation.

But since it is perfectly obvious giving a claim of the form “*P* because *q*” and identifying the *explanans* are acceptable forms of explanation, my strategy here is weakened by the need to resist these forms.

I have argued against [1b] and [1c] as general conceptions of the *nature* of the explanatory role of motivating reasons. In my view, a reason can play an explanatory role without being able to complete a ‘because’ claim or being the *explanans* of an action. I have also argued that it would be misguided to argue against this view by way of auxiliary constraints on explanation such as EXPLANANTIA ARE REASONS or TRANSLATABILITY. It would be a mistake to infer from any of this, however, that I think that these alternative ways of explaining action are illegitimate. Indeed, I happily allow that

A Φ d because *p*.

A is Φ ing because *A* believes that *p*.

What explains *A*’s Φ ing is that *p*.

The *explanans* of *A*’s Φ ing is that *p*.

are perfectly acceptable reason-giving and explanatory forms. I do not deny that reasons-explanations are often propositional, or that the explanation of an agent’s action often just is the content of the agent’s reason for acting given in propositional form. Neither is it my view that the only acceptable framework for giving reasons-explanations of action is that according to which we explain the action by giving the reason for which it was done. Instead, I only reject the thought that reasons-explanations *must* or *always* come in some or other of these forms; and I only insist that *one* legitimate way of explaining action is by giving the reason for which it was done.

So if by PROPOSITIONALISM we intend that in a reasons-explanation the *explanans* must be a proposition, and by that we mean that the *entity* that explains the action (whether by making it happen, causing it to come about, or some other such) must belong to the ontological category ‘proposition,’ PROPOSITIONALISM faces a steep challenge from the apparent variance in the nature of explanatory entities. It seems, in general, that states of affairs, events, mental states and other sorts of things can be explanatory entities. If we adjust PROPOSITIONALISM to mean that the explanatory entity must be capable of being expressed in propositional form, and by that we mean that it must be expressible in a that-clause, then we face difficulties with purposes and translatability in general. Even if we overcome these difficulties, there will be a serious question about why explanatory entities must *actually* be packaged in these ways, and whether anything deep about the nature of explanation is being revealed here. If we focus instead on the explanatory claims we offer (or the answers to “Why?” questions) with respect to actions, we might mean by PROPOSITIONALISM that such claims must be propositions. But this seems questionable in the light of fragmentary and infinitival explanatory clauses. It seems implausible and unmotivated to claim that only when we have something propositional in form do we have something acceptable *qua* explanatory claim. If again we adjust PROPOSITIONALISM to mean that what explanatory claims offer as explaining the action in question, we face the same difficulties.

In what sense is reasons-explanation propositional, then? It is not clear that there is anything both general and insightful to be said. There are some explanatory forms which are propositional; and there are some reason-giving forms which are propositional. If we opt to explain *A*’s Φ ing by way of a “*P* because *q*” claim, for example, then we commit ourselves to giving a propositional explanation of *A*’s Φ ing. If we opt to give *A*’s reason for Φ ing in propositional form, and also opt to explain *A*’s Φ ing by offering the

relevant reason-giving claim, then we also commit ourselves to giving a propositional explanation of A 's Φ ing. If we opt to say that A Φ d for the reason that r , and also to explain A 's Φ ing by stating that A Φ d for the reason that r , then we again commit ourselves to a propositional explanation of A 's Φ ing. We *may* opt for any of these methods, or others still; but we need not. We may opt instead to give A 's reason for Φ ing in infinitival form, and to explain A 's Φ ing by offering up the reason-giving claim. We may opt to explain A 's Φ ing by drawing attention, in any number of ways, to some state of affairs A found favorable. We could do similarly by casting something other than Φ ing (Ψ ing, for example), as the object of A 's desire. We *may* opt for any of these methods, or others still; but we need not. And so on. These considerations before us, I think we should say that the burden of argument is on those who insist on one or another of these particular explanatory forms at the expense of the others.

SUMMARY REMARKS

It has been a long and circuitous route to arrive at what I believe to be a somewhat simple and modest thought, which might serve as a summary of the main point of this chapter, and also as one of the main views that I am putting forward in this manuscript. The thought is just this. One way of explaining (in the sense delineated by reasons-explanations) an action is to give, specify, or otherwise reveal the reason for which the action was done. That being so, we can explain an action by offering up any claim or clause that gives, specifies, or otherwise reveals the agent's reason for acting. Since there are many adequate forms of reason-giving claims, there are many adequate forms of reasons-explanations. The legitimacy of these variations casts serious doubt on many supposed constraints on action explanation tacitly assumed by those theorizing about error cases. To the extent that arguments from error — whether aimed at telling us

what an agent's reason could or could not have been, or must have been — rely on these constraints, the wide variation in the acceptable forms of reason-giving claims also casts serious doubt on these arguments.

Breaking the Rules of Explanation

In the first two chapters of this manuscript, I defended a particular package of views from various objections. That package includes a view about the factivity of acting for a reason, which I have called the

Nonfactive View_M: it is not the case that in order for A to Φ for the reason that r , it must be the case that r .

It also includes a view about the cognitive constraint on acting for a reason, which I have called the

Belief View: in order for A to Φ for the reason that r , A must believe that r .

Finally, the package includes a view about the ontology of motivating reasons, which I have called

Non-Psychologism: it is not the case that whenever A Φ s for the reason that r , r is a psychological fact about A .

Much of this chapter has been spent excavating, elucidating, and refusing supposed rules of explanation. Once we see that we are free to break those rules (or, if you like, that they

are not genuine rules), I have suggested, we may begin to see how difficult it can become to argue convincingly against this package of views in certain ways. The Argument from the Factivity of Explanation, Argument from the Factivity of Acting for a Reason, and other arguments from error become much less forceful. We should not be persuaded away from the view that Edmund skates near the edge of the pond for the (false) reason that the ice in the middle of the pond is thin on the grounds that this falsehood cannot be the explanation of Edmund's skating, for it is not true that EXPLANANTIA ARE REASONS. Neither should we be persuaded, by similar reasoning, toward the view that Edmund's reason must have been a psychological state, such as his belief that the ice in the middle of the pond is thin. We might say that Edmund's reason must have been something he believed; and we should not be persuaded that it must have been something he knew by the thought that only then could we suppose that there is a reliable connection between Edmund's reason for acting and his acting for that reason. For our understanding of action is not constrained by any thought like NO HAPPY ACCIDENTS.

On the contrary, there is an embarrassment of explanatory riches at the disposal of such a package of views. That is in no small part due, I think, to the fact that reason-giving claims are subject to few, if any, formal or universal rules. So long as we can explain an action by giving the reason for which it was done, action explanation will be in this sense *permissive*: there are many permissible ways to explain actions, and many permissible forms of action explanations. On this view, there are many acceptable frameworks for giving reasons-explanations of actions, and inside each framework many acceptable reasons-explanations of any given action. In a way, I should like to all at once agree with Anscombe, Dancy, Davidson, Hyman, Sandis, Schueler, and others:¹⁰²

¹⁰² See Anscombe [1957a: 9], Dancy [2014: 90], Davidson [1963: 685], Hyman [1999: 443] Sandis [2013: 36-37], and Schueler [2003: 56], respectively.

The answer I shall suggest is that [intentional actions] are the actions to which a certain sense of the question ‘Why?’ is given application; the sense is of course that in which the answer, if positive, gives a reason for acting.

We can say that what explains the action is that it was done for the reason that p , without committing ourselves to saying that what explains the action is that p . It would remain true, however, that we explain an action by giving the reason for which it is done.

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

Reasons can be stated or given; and the canonical form of a sentence stating or giving a person’s reason for doing or believing something is ‘ $A \Phi$ ed because p ’.

What we infer from the above-mentioned explanatory statement, however, is not ‘that p ’, but that the agent thought or supposed that p ... what explains the action is the whole statement. More accurately, it is people who explain actions by *citing* one or more agential reasons, thereby implying strictly that (a) the agent took p and/or q to count in favour of her action and (b) acted accordingly.

Explanations of actions in terms of the agent’s reasons, I want to claim, work by citing the purpose or purposes for which the person who performed the action actually acted.

It is true, I think, that reasons can be answers to questions of the form “Why did $A \Phi$?”, that we can give reasons-explanations without giving the reason as the *explanans*, that reasons can rationalize actions, that we can give reasons-explanations in the form “ P because q ,” that we can give reasons-explanations in the form “ $A \Phi$ d in order to Ψ ,” and so on. So, although the tone of this chapter has perhaps been negative, ultimately I would like to endorse a framework for action explanation that is quite ecumenical. What I think is *not* true is that the success of any of these frameworks shows any of the others to be wholly unsuccessful. For, as Dancy and Davidson remarked,

There is more than one way of explaining an action... The [factive] form we have given our explanation commits us to the agent’s being correct... in supposing what his reason was to be the case. But we have other ways of explaining the action which

do not commit us in this sort of way, and which work equally well whether the agent was correct in this way or not.

To describe an event in terms of its cause is not to identify the event with its cause, nor does explanation by redescription exclude causal explanation.¹⁰³

There is indeed more than one way of explaining an action, and the goodness of one reasons-explanation of an action need not preclude the goodness of any other.

Before concluding, allow me to illustrate how I think this permissive approach toward reasons-explanations can afford great flexibility to a view packaging the Nonfactive View_M, the Belief View, and Non-Psychologism. We can return to our original case of Edmund, who takes it that THIN ICE: the ice in the middle of the pond is thin. So, when he skates, he keeps to the edge of the pond. In fact, the ice in the middle of the pond is not thin. If the Nonfactive View_M is correct, we can say that Edmund skates near the edge of the pond for the reason that THIN ICE. Since we can explain an action by giving the reason for which it was done, we give a reasons-explanation when we say that Edmund skates for the reason that THIN ICE. Can we, on this occasion, say that Edmund so skates *because* THIN ICE? Not unless we gloss ‘because’ here (à la Hornsby) with ‘for the reason’ — for otherwise we violate our commitment to FACTIVE EXPLANATIONS. Does this somehow indict the view that Edmund skates for the reason that THIN ICE? Not unless we subscribe to some implausible explanatory constraint such as NO FALSE THAT-CLAUSES, EXPLANANTIA ARE REASONS, or TRANSLATABILITY. Indeed, if we understand that a consideration can explain without being the *explanans*, we can even allow in some sense that “What explains Edmund’s skating near the edge is the consideration that the ice in the middle of the pond is thin.” The sense in which we allow this claim is revealed by the sense in which the consideration that THIN ICE can explain Edmund’s action: it can explain by being the consideration that motivated him to so act.

¹⁰³ See Dancy [2014: 346-347] and Davidson [1963: 695], respectively.

Now, if the Belief View is true, we can explain Edmund's skating by giving THIN ICE as his reason for acting, so long as he believes that THIN ICE. While this is at odds with the Psychologistic view that his reason is thin ice*: *he believes* that the ice in the middle is thin, this metaphysical fact does not preclude a Non-Psychologism from giving as a reasons-explanation the claim that "Edmund is skating near the edge because THIN ICE*. For there is no plausible rule that prevents us from picking out Edmund's reason by specifying it as the salient consideration which Edmund believed to be the case. Neither is there any plausible constraint that rules out specifying Edmund's reason in an infinitival clause, such as "in order to avoid thin ice." So, we could say that he skates as he does in order to avoid thin ice. Does this impugn the Belief View, since "in order to avoid thin ice" is not something capable of being believed to be the case? Not unless we insist on some implausible constraint like PROPOSITIONALISM. For otherwise, we can content ourselves with the thought that only if we opt to cast Edmund's action as a response to the consideration *that r* should we have to say that Edmund believes that *r*; and there is no such constraint on reasons-explanations such as "He did it in order to avoid thin ice." We could go on here; but I hope I have given some feel for the kind of flexibility that a view like the one I have proposed possesses.

Theorizing Without Arguments from Error

Allow me to tie all of this together in the following way. Arguments from factual error are supposed to present a general problem in action theory. They are supposed to show us that no agent can act for a reason that is not the case. According to some, then, these cases are supposed to show us that some kind of Factive View_M must be right. They are supposed to substantiate a general reason-robbing inference that no falsehood could be an agent's reason for acting. This leaves us with a question, though: what *is* the agent's

reason for acting in such cases? Psychologism, according to many, is the most appealing answer to this question. It is supposed to provide us with an appealing reason-replacing inference, to the conclusion that Edmund's reason, for example, must have been some psychological fact about him. Arguments from procedural error are supposed to present a problem even for Psychologistic views of acting for a reason. They are supposed to substantiate a reason-robbing inference that nothing not known by the agent could be the agent's reason for acting. But again, this leaves us with the question of what the agent's reason *is* in these cases. The Knowledge View, according to some, is the most appealing answer to this question.

So one can see how central the analysis of error cases has been to action theory in general. But I have argued that, quite generally, these arguments are misguided; and if we adopt the kind of permissivism about reason-giving claims and reasons-explanations that I have favored, then a theory comprised of the Nonfactive View_M, Belief View, and Non-Psychologism is not easily damaged by such criticisms. If I am right about that, then, one may be left wondering how theorizing about acting for a reason is to proceed. If what I have said is correct, then one might think that error cases are unlikely to tell one way or the other about the main views in the area. But it might also be thought that the plausibility of theories of acting for a reason is heavily dependent upon their being able to say the right things about error cases.

I do think that, unless the relevant explanatory constraints can be made plausible, or the argument from error can somehow be worked out without them, we should be much more cautious about the use of error cases in action theory. I also hope that what I have argued for here encourages some fairly serious re-evaluation of the theoretical landscape. But I think it would be an overreaction to say that any of this presents an impasse in our theorizing, or that it places the main competitors necessarily at a

stalemate. For recall that all of the arguments in question have made heavy use of the explanatory role of reasons for acting. What unifies these arguments, both in their apparent force and their ultimate failure, is the kinds of claims they make connecting reasons and explanations. If I am right, then we theorists should reconsider both our preconceptions about the nature of action explanation as well as its implications for action theory in general. For example, the precise sense in which reasons for acting play an explanatory role certainly deserves more attention; and action theory would do well to engage in debate, as philosophers of science have, about what it means for explanation to be factive — or whether explanation is factive at all.

More significantly, though, we should keep in mind that the explanatory connection between an action done for a reason and the reasons for which it the action is done is only one feature among very many that we might look at when evaluating theories in the area. Often underemphasized in these debates is, for example, plain conceptual analysis of *acting for a reason* and *reason for acting*. Often forgotten is the deep apparent similarity between acting for a reason and acting for a purpose (I should have liked to say much more about this). Often completely ignored is the tempting thought that whatever is said about action in these debates should be transplantable into theories of belief, believing for a reason, and reasons for believing (I shall say more about this in the next chapter). If the thought that motivating reasons for acting are propositions cannot be supported, for example, certain views in epistemology might be found implausible; on the other hand, if propositionalism about reasons for belief is undeniable, this may be reason to accept it for action. Looking elsewhere: very little has been said about what it is to respond to a reason; but certainly acting for a reason is a way of responding to a reason. Or perhaps we could, following Anscombe, look to the ways in which the nature of intention and intentional action might constrain what can plausibly be said about reasons

and acting for a reason. For example, some have argued for cognitive constraints on intention and intentional action; but if acting for a reason is a species of intentional action, then one might expect this to affect what we say about cognitive constraints on acting for a reason. Finally: not only might each of these issues have telling implications for acting for a reason, but also vice versa. For example, it is worth noticing that the factivity of *normative* reasons is almost always taken for granted. But when it is argued for, it is argued for in ways quite similar to the arguments criticized in this paper. Perhaps this issue also deserves some more attention.

Chapter Four: Motivating Reason to Slow the Factive Turn in Epistemology

In this final chapter, I would like to spend some time addressing some aspects of those additional questions just mentioned. I hope that the discussion so far has revealed that, though the package of views I have defended would be a minority view, practical philosophy has been increasingly considering the possibility of such nonstandard views (such as the Nonfactive View_M, the Belief View, and Non-Psychologism). On the other hand, views analogous to the Factive View_M, Knowledge View, and Psychologism have been gaining favor in epistemology. This makes for an intriguing time to be thinking about reasons in general. In this chapter, I hope to lay the foundation for a project that encourages more exchange between theorists of practical and epistemic reasons. I hope to do so by illustrating how such an exchange would be fruitful vis-à-vis the issue of factivity. The primary aim of this chapter will be to slow the ‘factive turn’ in epistemology, as it were, by appealing to some of the work done in the first three chapters on practical reasons. This primary aim is a part of a larger aim to provide a unified theory of reasons: motivating, normative, explanatory, practical, epistemic, moral, and otherwise. But the endgame of this chapter will be limited to a view according to which *normative* reasons for believing (or *evidence*, if you like) are, as I have argued concerning motivating reasons for action, not factive. I will argue that if motivating reasons for acting are not factive, then normative reasons for believing are not factive. In conjunction with the case that I have made for the nonfactivity of motivating reasons, this conditional helps to secure the conclusion that normative reasons for believing are not factive. But even without the case I have made for the antecedent of this conditional, I hope to be revealing a connection between motivating reasons for acting and normative reasons for believing that is interesting in its own right. Indeed, I suspect that the implication of the main argument

of this paper will strike many as surprising, since it suggests that, in order to successfully establish the view that normative epistemic reasons are facts, one must fend off the view that motivating practical reasons need not be facts.

The course of the chapter is as follows. I begin in the first section with some preliminary remarks about two standard distinctions in the theory of reasons: that between motivating and normative reasons, and that between practical and epistemic reasons. I also briefly review and spell out what the Factive and Nonfactive View of each of these looks like. I proceed in the second section of the paper to lay out an important claim on which my argument relies: the claim that our positions on the factivity of epistemic and practical *normative* reasons should agree. I defend this claim of uniformity from an initial concern, namely, that epistemic reasons owe their distinctive character precisely to their relation to the truth. That being the case, one might think that we have a special reason to think that epistemic reasons are facts — one which does not also lend support to the idea that practical reasons are facts. But I show that this concern is misguided.

With this claim of uniformity on board, I spend the next two sections of the paper establishing the following claim: if motivating reasons need not be facts, then normative reasons need not be facts. I begin in the third section with a brief reminder as to why one might think that motivating reasons need not be facts. I then move to motivate the thought that in some cases where A Φ s for the reason that r and it is not the case that r , A 's Φ ing is rational; not only can an agent act in the light of a falsehood, but an agent can also act *rationally* in the light of a falsehood. I also claim that a plausible conception of rationality is one on which A 's Φ ing can be rational only if A 's Φ ing is done for some good reason. If this is all correct, then some false reasons for which agents act must also be good reasons — and so some normative reasons must not be facts.

I connect all of the dots in the fourth section by briefly laying out my main argument against the factive turn in epistemology. In conjunction with the claim of uniformity from the second section, the previous two sections give us what we need to mount a somewhat surprising argument against factive views of epistemic normative reasons. In the fifth section I address two objections, both of which try to thread the needle a bit, by forcing a wedge between rationality and reasons-responsiveness. According to the first of these, since in the cases of interest A has no good reason to Φ , we must stop short of attributing *rationality* to A 's Φ ing. According to the second of these, we might concede that in these cases A 's Φ ing is *rational*, but not in any sense that implies that A Φ s for a good reason. In the sixth section I consider a somewhat different way of resisting the central argument of the chapter. One might grant the core argument of my paper — that the factivity of motivating and normative reasons rise and fall together — and simply use it as a *reductio* against the view that practical motivating reasons need not be facts. I suggest that this is not as easily done as one might initially suppose. I conclude in the seventh section by looking forward, sketching out what I call the practical theory of epistemic reasons. This is a theory which takes the case for the Nonfactive View_M, the Belief View, and Non-Psychologism seriously, and takes up analogous positions with respect to reasons for belief. In giving this sketch I hope to be pointing toward that larger project, alluded to earlier, of providing a comprehensive and unified theory of reasons.

PRELIMINARY REMARKS

As I said, I am primarily concerned to establish the claim that: if motivating reasons need not be facts, then normative reasons need not be facts. I have also said that my demonstration of this claim will operate primarily on “practical” reasons, but with

clear implications for “epistemic” reasons. Such a project calls for clarifications on two fronts.

First, there is the standard distinction between motivating reasons and normative reasons, which we already have on board. We have seen that there is a debate concerning each of these as to whether they are facts. We have so far discussed the first — the debate between the Factive View_M and the Nonfactive View_M — in great detail. We noted at the beginning of the first chapter that this should be distinguished from the debate between the

Factive View_N : in order for r to be a reason for A to Φ , it must be the case that r and the Nonfactive View_N (which is the denial of the Factive View). To borrow the earlier example: even if you know that I am running through the halls for the reason that I am late to give my eleven o’clock lecture, you might wonder whether there is any *good* reason for me to be running through the halls. Indeed, you might wonder whether I am, in running through the halls, acting for a good reason. That is: is the consideration which motivates me to run through the halls (*that I am late*) also a consideration that favors, justifies, or makes right my running through the halls? According to the Factive View_N , the consideration that I am late can only favor my running through the halls if it is the case that I am late. The Nonfactive View_N denies this, holding that the consideration that I am late is capable of favoring my action even if it is not the case that I am late.

Here I have made the distinction between motivating and normative “practical” reasons, as they are called. But, as I have suggested, analogous issues manifest with respect to so-called “epistemic” reasons as well. This is the second front on which my project is in need of clarification. Borrowing from the literature is less useful in making

this distinction. For example, it is not uncommon to encounter the thought that practical reasons are reasons for acting (that is: considerations that favor actions) while epistemic reasons are reasons for believing (considerations that favor beliefs). But it is not difficult to see why that is not a sufficient way of making the distinction. The fact that my believing it to be so will make me happy certainly seems to be a reason for believing it to be so, but I doubt that any of us want to classify this consideration as an epistemic reason (at least, not in the standard sense, which is closely tied to the notion of *evidence*). It is no more promising to attempt making the distinction in terms of the *contents* of the considerations. In the example considered just above, the consideration that *my clock reads five past eleven* might be both a practical reason (in which case it probably favors my running through the halls) and an epistemic reason (in which case it probably favors my believing that I am late).

I shall try to make the distinction in a different way. In doing so I shall follow roughly the distinction as laid out by Joseph Raz.¹⁰⁴ The distinction between practical and epistemic reasons is most easily thought of as a distinction between two kinds of normative reasons. According to this framework, practical reasons are considerations that owe their normative force to their relation to values. Epistemic reasons, on the other hand, are considerations that owe their normative force to their relation to the truth (which may itself be a kind of value). My being late to give my eleven o'clock lecture seems to be a practical reason in this sense: it favors my running through the halls, and it so favors in virtue of its relation to certain values (such as that of being punctual). As such, this is a practical reason; but there are epistemic reasons nearby. Suppose that you ask me

¹⁰⁴ See Raz [2011: 36-58] for further explication of this distinction, which I take to be standard. Notice that this crucial distinction remains intact even on views according to which truth is a kind of value (and epistemic reasons can therefore be thought of as a kind of practical reason). The distinction would then be grounded in a distinction between two kinds of values. Thanks to an anonymous referee for encouraging me to address this possibility.

why I believe that I am late, and I respond that the clock in my office reads quarter-past eleven. As such, this fact about my clock seems to be an epistemic reason: it favors my believing that I am late, and it so favors in virtue of its relation to the truth (about what time it is).

In the literature on practical and moral reasons, the Factive View_N is so far in the majority that it is often assumed without argument. Instead, the debate tends to be about *in what sense* reasons for acting are facts, or *what kind* of facts reasons for acting are. Here is Maria Alvarez:

What sort of thing is a normative reason?... There is consensus that normative reasons are facts (Raz 1975; Scanlon 1998, though the consensus is not universal. The question is complicated by disagreement about what facts of any kind are: are they concrete or abstract entities? Is a fact the same as the corresponding true proposition, or is the fact the “truth-maker” of the proposition?¹⁰⁵

While there is controversy about the precise nature in which reasons for action are facts, there is little to no discussion as to *whether* they are facts. In the literature on epistemic reasons, things are a bit different. On the orthodox analysis of evidence as good reason to believe, the debate about the factivity of normative reasons in epistemology might be hiding under the label “the factivity of evidence.” On that topic, the Factive View_N seems only recently to be coming into some favor. It is bolstered by recent defenses of the view that one’s evidence consists in those things one knows. Here is Thomas Kelly, charting out some of the territory.

A second important division under this general heading is between theorists who hold that what evidence one has is completely fixed by one’s non-factive mental states and theorists who deny this. This division between ‘evidential internalists’ and ‘evidential externalists’ cuts across the propositionalist/non-propositionalist distinction. Thus, a proponent of the classic sense data theory mentioned above will hold that one’s evidence is non-propositional and entirely a matter of what

¹⁰⁵ See Alvarez [2016: 6-8]. For views representing this mainstream assumption, see Dancy [2000; 2004], Darwall [1983], Parfit [2011], Raz [1975], Scanlon [1998], and Smith [1994].

mental states one is in; this second, internalist commitment will be shared by various paradigmatic propositionalists, e.g., a coherentist who holds that one's evidence consists of token beliefs that are sufficiently well-integrated with one another. On the other hand, a propositionalist who thinks that one's evidence consists of all and only those propositions that one knows (e.g., Williamson 2000) will think that what evidence one has depends not only on what non-factive mental states one is in but also on how things stand in the external world. This 'externalism' about evidence will also be endorsed by a non-propositionalist who takes at face value the many ordinary assertions which suggest that physical objects can themselves count as evidence in certain contexts.¹⁰⁶

One can see here that there are a variety of ways to hold Factive View_N the concerning reasons for believing (or evidence). But they all hold, in one way or another, that such reasons are always facts.

UNIFORMITY

So much for preliminaries. I said at the outset that this paper is a part of a larger project that encourages more exchange between theories of practical and epistemic reasons. But, given the deep difference in the respective natures of practical and epistemic reasons, one might worry about the prospective fruitfulness of such an exchange. If there is legitimate reason to worry, the main argument of this paper might be robbed of some of its force. That is because the argument relies in part on the claim that I will call

UNIFORMITY: *ceteris paribus*, the best theory of reasons will hold just one position about the factivity of reasons, which is true of every kind of reason.

I mean this claim to apply within each of the realms of motivating and normative reasons independently, but across normative realms (moral, epistemic, practical, aesthetic, or

¹⁰⁶ See Thomas Kelly [2014: 5] for the standard thought that evidence is good reason to believe, and Kelly [2008: 941-943] for the quoted passage (and also the rest of the paper for a nice discussion of how the debate concerning the factivity of evidence has developed). No doubt the work of Williamson [2000] has been enormously important on this front. See also Littlejohn [2012; 2013a; 2013b].

otherwise) therein. Suppose, for example, that we become convinced that the Factive View_N of epistemic reasons is correct. Then, according to the principle of UNIFORMITY, there is theoretical pressure to endorse the Factive View_N of practical reasons. Similarly: if one becomes convinced that the Nonfactive View_M of practical reasons is correct, then one is under some theoretical pressure to endorse the Nonfactive View_M of epistemic reasons. For my part, I find this commitment to UNIFORMITY to be simply intuitive. It seems a good general method of inquiry to search out the most uniform and simple theory of the phenomena in question. Here is Maria Alvarez, expressing some sympathy for a similar thought applied specifically to reasons:

It is more plausible to assume, at least at the outset, that there is some continuity or common thread linking the use of the term ‘reason’ in these different contexts—an assumption that would remain tenable even if the roles that reasons can play in some contexts... should turn out to have some peculiar features not found elsewhere.¹⁰⁷

It is a further question, of course, whether the factivity (or nonfactivity) of reasons is one of these “common threads.” But I find no good reason not to extend the general preference for uniformity in theories to the subject matter of reasons. Is there any reason, for example, to expect that the considerations which motivate us to act and those which motivate us to believe should be different in this respect? We might hope instead that the phenomenon of rational motivation is uniform in nature. Since this principle is initially appealing, and I know of no good argument against it, I will treat it as a kind of working assumption.¹⁰⁸

¹⁰⁷ See Alvarez [2010: 8].

¹⁰⁸ One might think that the constraint could be supported by considerations of the theoretical virtue of simplicity. Maria Alvarez has convinced me that it is not straightforward how this could be done, so I have left out appeals to simplicity. Suppose we take a nonfactive view of motivating reasons for acting, for example. Presumably, we still admit facts into our ontology; so it is not clear that holding also that motivating reasons for believing are also not factive necessarily makes our ontology simpler.

Before moving on, I want to address one possible, though misguided, source of doubt about. The doubt may be due to the difference between practical and epistemic reasons just laid out. The distinguishing mark of epistemic reasons is supposed to be their relation to the truth (practical reasons stand in no such special relation); and the debate about the factivity of reasons is, essentially, a debate about whether something must be true in order to be a reason. Should we not, then, expect a deep difference between practical and epistemic reasons on the issue of factivity? I think not. There are two questions that need to be separated here. One is the question of what makes something a reason *at all*; another question is what makes something a reason *of a particular kind*. Put another way: one question is whether some consideration must be true (or be the case) in order to favor something else, while another question is what explains or grounds the ability of those things that do favor to favor in the way that they do.¹⁰⁹ The distinction just laid out, of course, gives a standard answer to the second question. But that answer does not imply any answer to the first question. Suppose that r is an epistemic reason for some agent A to Φ (where Φ ing is, say, believing that p). The standard distinction tells us that what makes r an epistemic reason to Φ , rather than a practical reason to Φ , is the relation that r bears to the truth of p (for example, r might make it more likely to be true that p). This, of course, says nothing about whether r *itself* is true; it is not the truth of r , but rather the effect that r has on the truth of p , that is important for taking r to be an epistemic reason. But the debate over the factivity of reasons, applied here, would be a debate about whether r *itself* must be true in order to be a reason. Put yet another way: the standard distinction tells us that there are two ways in which a consideration can be made normatively relevant: one is via relation to values, and another is via relation to the truth.

¹⁰⁹ Thanks to Jonathan Dancy here.

Being a fact is neither of these, and indeed hardly seems to be a way of being normatively relevant at all.

So we can see clearly that the special relation that epistemic reasons bear to the truth is really no reason at all to suspect that they will be deeply different from practical reasons vis-à-vis the issue of factivity. If the principle of UNIFORMITY is correct, there may be serious implications for theorizing about reasons — especially if the natures of motivating and normative reasons can be connected in the manner I shall go on to claim in this paper. Together, these two claims amount to the thought that theories of reasons must take just one position on factivity, which holds for all kinds of reasons: motivating, normative, practical, epistemic, and perhaps otherwise. Consequently, someone who sets out to establish the factivity of normative epistemic reasons, for example, will be saddled with defending the factivity of these other varieties of reasons as well.

ACTING RATIONALLY IN THE LIGHT OF A FALSEHOOD

With these initial remarks in mind, I move in this section toward my central claim that if motivating reasons need not be facts, then normative reasons need not be facts (or, in other words, if the Nonfactive View_M is correct, then the Nonfactive View_N is also correct). I will do so in part by working through a series of cases. We can begin with a familiar sort of case.

ONE SKATER: Imagine a pond that has thin ice in the middle. Edna takes it that the ice in the middle of the pond is thin. So, when she skates, Edna keeps to the edge of the pond. You are on a nearby hill, and you see Edna skating. You ask her why she is skating as she is, and she tells you that the ice in the middle of the pond is thin.

We saw in the first chapter that, both intuitively and theoretically, it is plausible to suppose that Edna is skating near the edge of the pond for the reason that THIN ICE: the ice in the middle of the pond is thin. We can next consider the case of

TWO SKATERS: Imagine two adjacent ponds. Edna takes it that the ice in the middle of one pond is thin. So, when she skates on it, she keeps to the edge of the pond. Edmund takes it that the ice in the middle of the other pond is thin. So, when he skates on it, he keeps to the edge of the pond. You are on a nearby hill, and you see both skaters skating — but you have no view about how things are with the ice. As they finish, you ask them both why they keep to the edge while they skate. They both tell you that the ice in the middle of the pond is thin. When you get home, your sister informs you that the ice in the middle of Edmund's pond is just fine.

As I have argued, we should understand both characters in this case in just the same way as we understand Edna in ONE SKATER. In the case, you would readily accept both responses to your “Why?” question. Both Edna and Edmund have presented you with the consideration that motivated them to skate near the edge of the pond, and you can use this consideration to make sense of their skating near the edge of the pond. The information your sister gives you when you get home makes no difference to any of this. What your sister tells you about Edmund's pond does not leave you confused as to what Edmund was doing or why he was doing it (although it might leave you confused as to why Edmund *believes* what he does). Despite the fact that Edmund was wrong about the

way the world is, you can still understand his behavior by a thought such as, “His reason for keeping to the edge was that the ice in the middle of the pond is thin.”

At this juncture, we can set aside whether I have been successful in establishing that one can act in the light of a falsehood. As has been noted, we theorists might not be interested only in what the reasons for which agents act are or could be. We might also be interested in whether, when an agent acts for some reason or another, the agent is acting *rationally* in so doing. We might allow that, for example, Edmund acts for the reason that THIN ICE even though it is not the case that THIN ICE. But when he so acts, is his action rational?

To get at this question, consider a second pair of cases.

JUST SKATING: Imagine that Edmund is about to go skating. As he heads out, Edna tells him that the ice in the middle of the pond is thin. When Edmund goes out to the pond, the ice in the middle of the pond looks thin to him. Edmund takes it that the ice in the middle of the pond is thin. So, when he skates, Edmund keeps to the edge of the pond. Meanwhile, you are on a nearby hill, and you see Edmund skating. You approach him and ask him why he is skating as he is. He responds by saying that the ice in the middle of the pond is thin.

There is nothing unusual about this case. This is the kind of situation that any of us living near a pond might find ourselves in on a nice winter day. I think that any of us in this case would accept Edmund’s response to our inquiry. If it occurred to us to ask the question at all (which, tellingly, I think it would not), we would deem Edmund’s action to be rational. I also think we would be right on both counts: Edmund *does* act for the reason

that THIN ICE, and Edmund's skating near the edge, in light of the consideration that THIN ICE, *is* rational. Suppose though, that I add one detail to the case:

IGNORANT SKATING: Imagine that Edmund is about to go skating. As he heads out, Edna tells him that the ice in the middle of the pond is thin. When Edmund goes out to the pond, the ice in the middle of the pond looks thin to him. Edmund takes it that the ice in the middle of the pond is thin. So, when he skates, Edmund keeps to the edge of the pond. Meanwhile, you are on a nearby hill, and you see Edmund skating. You approach him and ask him why he is skating as he is. He responds by telling you that the ice in the middle of the pond is thin. Unbeknownst to Edna, Edmund, and you, the ice in the middle of the pond is not thin. When you get home after skating, your sister tells you that the ice in the middle of the pond is not thin.

Now, it seems to me that this second pair of cases is dialectically analogous to the first pair of cases, in the following way. We have the intuition about JUST SKATING that Edmund's skating is rational. Just as with the transition from ONE SKATER to TWO SKATERS, adding the last detail here does nothing to alter our initial judgment. In moving from JUST SKATING to IGNORANT SKATING, Edmund goes from having a true belief about the pond to having a false belief; but so far as the rationality of his action is concerned, nothing has changed. Even when you learn that you, Edna, and Edmund were mistaken about whether the ice is thin, you still (correctly, I think) judge Edmund's skating near the edge to be rational.

Some may be tempted to deny what I have just said. That is, some will want to uphold the (intuitive) claim that Edmund's action is rational in JUST SKATING, but deny it

in IGNORANT SKATING. But doing this requires not just denying the intuitive verdict in such cases, but also denying plausible general claim about rationality. This is the claim that

ERROR NEED NOT CREATE IRRATIONALITY: in many cases, the difference between A 's being correct or incorrect about the truth of some consideration p is not the kind of difference that can, on its own, make a difference to whether A is Φ ing rationally.¹¹⁰

There are some facts such that an agent's being wrong about them pretty clearly does impugn that agent's rationality. If Edmund thought that there being five clouds in the sky was reason to keep to the edge of the pond, for example, and this was the only consideration that he could mention in giving an account of his skating, we might have our doubts about the rationality of his action. Perhaps we might say something similar about Anscombe's well known example:

If someone hunted out all the green books in his house and spread them out carefully on the roof, and gave one of these answers to the question 'Why?' his words would be unintelligible unless as joking and mystification. They would be unintelligible, not because one did not know what *they* meant, but because one could not make out what the man meant by saying them here... If we say 'it does not make sense for this man to say he did this for no particular reason' we are not 'excluding a form of words from the language'; we are saying 'we cannot understand such a man'.¹¹¹

We could go one further than Anscombe here, and say that we cannot consider this man's action to be rational. If this man answered our "Why?" question with something like, "So that the mailman will come sooner," we would judge his action to be positively

¹¹⁰ Thanks to an anonymous referee here, whose comments helped me to formulate the claim in this way.

¹¹¹ See Anscombe [1957: 26-27].

irrational. But Edmund's being wrong about whether THIN ICE does not seem to fit this bill — especially in cases like IGNORANT SKATING, where Edmund seemingly has good reason to believe that THIN ICE, and there is a clear connection between the consideration that THIN ICE and the act of skating near the edge of the pond. Deeming his action irrational, it seems to me, would be at odds with the plausible thought that ERROR NEED NOT CREATE IRRATIONALITY.

In the light of these considerations, I think that the plausible thing to say is that Edmund's skating near the edge is rational — even IN IGNORANT SKATING. It seems possible not only to act in the light of a falsehood, but also to act rationally in the light of a falsehood. But what does this have to do with the factivity of normative reasons? To make the connection, consider one final case:

SENSELESS SKATING: Imagine that Edmund is about to go skating. As he heads out, Edna tells him that the weather is nice, and that it is a great day for skating. When Edmund goes out to the pond, he does not notice anything strange about the ice. So, when he skates, Edmund keeps to the edge of the pond. Meanwhile, you are on a nearby hill, and you see Edmund skating. You approach him and ask why he is skating as he is. He responds by telling you that the weather is nice, that it is a great day for skating, and that he noticed nothing strange about the ice.

Now, I think that any of us in this case would have our doubts as to the rationality of Edmund's action. We may find it difficult to even make his action *intelligible* (as indication of this, notice the difficulty encountered in interpreting the “So” in the middle of the story). I think that the most plausible explanation of these reactions is that we cannot find, anywhere in Edmund's account of his action, any good reason for doing what he is doing.

Indeed, we cannot find in his account anything we could understand him as *taking to be* good reason for doing what he is doing (despite the fact that he has offered up some considerations in answering our “Why?” question). This makes it hard to suppose that his action is rational; for an action to be rational, it is natural to think, it must be done for some good reason.

These observations suggest another way of making the point here. Even independent of our judgments about cases like SENSELESS SKATING, one plausible (even if naive) way of conceiving of rationality is as having an intimate connection to good reasons — considerations that favor responses. We naturally conceive of rationality as some kind of appropriate reasons-responsiveness. If acting rationally is, *inter alia*, responding appropriately to reasons, then an action must at least be an appropriate response to some consideration in order to be rational. While this is a standard way of thinking about the rationality of belief, it is somewhat less common to find the thought clearly expressed about the rationality of action. But here are Juan Comesaña and Matthew McGrath expressing this intuitive conception:

One does something rationally only if one has reasons that make it reasonable for one to do it and one does it on the basis of some (sub-)set of those reasons, i.e., one does it “for” those reasons.¹¹²

I find this to be a very appealing thought. There are some who try to distinguish between rationality and reasons-responsiveness, and I will address that way of thinking momentarily.¹¹³ Notice, though, that my claim here (and what my argument will require) is not that rationality *just is* responding appropriately to reasons. Instead, I claim only that appropriate reasons-responsiveness is necessary for rational action, so that if one acts rationally, it follows that one responded appropriately to some reason or another.

¹¹² See Comesaña and McGrath [2014: 62].

¹¹³ Thanks to Veli Mitova for pointing out the relevance of such views.

Something like this conception of rationality would explain why Edmund's action in JUST SKATING and IGNORANT SKATING seems rational, but his action in SENSELESS SKATING does not. His action in the first two cases seems to exhibit appropriate responsiveness to the consideration that THIN ICE, but his action in SENSELESS SKATING shows no signs of responding appropriately to *any* considerations.

MOTIVATING REASON TO SLOW THE FACTIVE TURN

We now have in place the materials for my central argument against the Factive View_N of epistemic reasons. It goes as follows.

- [1] In some cases where A Φ s for the reason that r , it is not the case that r .
- [2] In some cases as described in [1], A 's Φ ing is rational.
- [3] A 's Φ ing can be rational only if A Φ s for a good reason.
- [4] In such cases as described in [2], r must be a good reason for Φ ing.
- [5] Some good reasons for Φ ing are things that are not the case.
- [6] If the Nonfactive View_N is true of practical reasons, then it is true of epistemic reasons.
- [7] Some epistemic normative reasons are things that are not the case.

I argued extensively for premise [1] in the first chapter. Premise [2] is supported by judgments about Edmund's action in JUST SKATING and IGNORANT SKATING, as well as the plausible principle that ERROR NEED NOT CREATE IRRATIONALITY. Premise [3] can be provided by a plausible explanation of our confusion about Edmund's action in SENSELESS SKATING, as well as the plausible thought that rationality is some kind of appropriate reasons-responsiveness. Claim [4] follows from premises [1]-[3], on the

simplifying assumption that, for example, THIN ICE is the only consideration Edmund is responding to in IGNORANT SKATING. Claim [5] follows from claim [4]. Premise [6] is an application of the principle of UNIFORMITY discussed in section two. The conclusion, [7], follows from claims [5] and [6]: epistemic reasons for belief — or evidence, if you like — need not be facts.

The philosophical value of this argument, I hope you will agree, is not limited to its presenting a strong challenge to the Factive View_N of epistemic reasons. That particular challenge relies on the truth of claim [1]. But, if the reasoning after claim [1] is sound, then we will have at least shown that if the Nonfactive View_M of practical reasons is true, then the Nonfactive View_N of epistemic reasons is true. If the argument is valid, this conditional claim will stand even if claim [1] is false or implausible. As I said at the beginning of this chapter, I think this would demonstrate an important and surprising connection between practical motivation and epistemic rationality. If this argument is correct, those wishing to force the factive turn in epistemology, as it were, will be saddled with defending the Factive View_M of practical reasons.

REASONS AND RATIONALITY

Given the discussion of the first chapter, we might predict that many will be tempted to save the Factive View_N of epistemic reasons by rejecting premise [1]. Since we have treated that issue extensively already, I move to treat two other serious objections to this argument. These two objections, if successful, would serve to sever the alleged connection between the Nonfactive View_M of practical reasons and the Factive View_N of epistemic reasons.

Appropriate Reasons-Responsiveness Without Rationality?

The first of these objections is most straightforwardly thought of as an objection to premise [2]. According to this objection, although there might be something praiseworthy about what Edmund is doing in cases like TWO SKATERS and IGNORANT SKATING, we should stop short of saying that his action is *rational*. In order for his action to be rational, it must be a response to some good reason for acting. But in such cases, there are no good reasons for Edmund to do what he is doing. We know this because there is no feature of the world that favors his doing what he is doing. There is a feature that *Edmund might think* favors his skating near the edge, but alas, that feature is not a real feature of the situation. Since rationality consists in some kind of appropriate reasons-responsiveness, we must stop short of saying that his action is rational.

One piece of this line of objection I will put off until later, since it is relevant to the other lines of objection as well. At this time I will make two points of response. My first response is to ask those hesitant to ascribe rationality to Edmund's action to consider what terms of appraisal they might find applicable in its stead. We might naturally say, for example, that his action is *reasonable*; that his action is perfectly *sensible*; his action is easily *intelligible*; his action is *understandable*; it is *what any rational agent in his position would do*; and so on. Indeed, it might seem that skating near the edge of the pond is the *only* thing Edmund could have done such that we would describe his action in all of these ways. Suppose, for instance, that Edmund, taking it that THIN ICE, proceeded to skate incautiously all over the pond. Here I take it that we would all be well disposed to rationally criticize Edmund's action (unless some detail were added to the story: perhaps Edmund is an adventurous fellow looking for a thrill, perhaps his friends put him up to it, perhaps he first skates timidly as a way of investigating whether the ice really is thin, et

cetera).¹¹⁴ Such as things are in TWO SKATERS and IGNORANT SKATING, though, I think that we would make any of the positive appraisals of Edmund's action without hesitation — and we would be correct in making those appraisals. But with this cluster of concepts before me I now feel puzzled about the initial hesitation to attribute rationality to Edmund's action. It would be a fairly foreign, and perhaps somewhat artificial notion of rationality such that all of these things can truly be said of Edmund's action, and yet we cannot truly describe his action as rational. So, to anyone offering this sort of resistance, I would first invite them to provide and motivate that further conception of rationality.¹¹⁵

The second point of response to this line of objection is that while Edmund's action may be ill-informed, it seems to me very difficult to uphold the claim that it is *irrational*, or that what he is doing is *rationaly impermissible*. This is partly due to the plain intuitive judgment that Edmund's action is not irrational. But this is also partly due to theoretical difficulties that arise if we render this verdict about what Edmund is doing. If Edmund's action is irrational or rationally impermissible, there must be something that makes it so; it will not just be a brute fact that it is rationally impermissible for Edmund to skate near the edge of the pond. It seems generally plausible that

NO BRUTE (IR)RATIONALITIES: If A 's Φ ing is rational (or irrational), there must be some Ψ in virtue of which A 's Φ ing is rational (or irrational).

¹¹⁴ See Williams [1981: 109] for a similar thought — although I do not want to imply that I am sympathetic to the sort of wholesale internalism represented there.

¹¹⁵ I do not want to create the impression that this is an impossible task. Littlejohn [2017], for example, has tried his hand at teasing some of these notions apart (thanks to an anonymous referee for bringing this to my attention). Adjudicating whether such projects are successful is beyond the scope of this essay, but I say a bit more in the sixth section.

But what feature of the situation could make it rationally impermissible for Edmund to skate near the edge of the pond? If we are developing this line of objection with a sympathetic eye toward the Factive View_N, perhaps we might try to say that it is the fact that it is not the case that THIN ICE which makes Edmund's action irrational. But surely that is not correct. The mere fact that it is not the case that the ice in the middle of the pond is thin cannot make it positively *irrational* to skate near the edge of the pond (though it might be thought to bear on the rationality of Edmund's *believing* that THIN ICE).¹¹⁶ At most, we might say (in line with the Factive View_N) that its not being the case that THIN ICE prevents THIN ICE from making any action rational. With that in mind, perhaps the strategy might be to say, in view of NO BRUTE (IR)RATIONALITIES, that just as Edmund's skating cannot easily be deemed irrational, neither can it easily be deemed rational. There is simply no feature of the situation that can be adduced to adjudicate the issue in either direction. There is an issue about whether some question is being begged in favor of the Factive View_N here. But I think that issue can be sidestepped in light of a more general point. Currently, we are considering the rational status of Edmund's skating near the edge of the pond, limiting ourselves to two choices: either his skating is rationally permissible, or it is not. Perhaps the same argument that might underpin the claim that it is permissible can also be used to argue that it is not permissible. But even in that case, it seems that we are left with the bare intuition for making our choice here; and the intuitive judgment about the case is surely that Edmund's action is rational.

One might think that the trouble here is being created by a false dilemma, in that we are wrongly supposing that we must make a course-grained, conclusive judgement about whether Edmund's action is all-things-considered rational (or not rational). Perhaps

¹¹⁶ This might be thought; but as I will soon suggest, this thought is no more plausible than the thought presently under consideration.

instead we should admit the appealing idea that rationality comes in degrees. Perhaps Edmund's action is rational *to some degree*, but not to the degree such that we should judge it to be rational *full stop*. I do not think this suggestion can help the one trying to plausibly avoid the verdict that Edmund's skating near the edge of the pond is rational. For the principle NO BRUTE (IR)RATIONALITIES is no less plausible when we think of rationality as a gradable attribute than it was when we were thinking of rationality in more course-grained terms. If Edmund's skating near the edge of the pond has *any* rationality to recommend to it, but not as much as it otherwise could, there must be some feature(s) of the situation in virtue of which that is so. But in virtue of what could we award Edmund's skating near the edge any degree of rationality, if not that it is the appropriate response to the consideration that THIN ICE? But then if we are prepared to deem his action rational (to some extent) in virtue of *that* fact, in virtue of what do we temper our estimation of the rationality of his action? We might say (again in line with the Factive View_N) that the fact that it is not the case that THIN ICE is what prevents us from judging his action to be more fully rational. But this does not seem to be the right kind of thing to make this kind of difference. As John McDowell observes,

It does not add to the explanatory power of explanations given by using those forms if things are as the agent relevantly takes them to be. If things are that way, that is just an extra fact about the situation; the action would have had the same intelligibility... if the act had been wrong in taking things to be that way.¹¹⁷

McDowell's remark concerns the rational *explanation* of action; but I think an analogous remark applies to the rational *evaluation* of action. If we say that in skating near the edge of the pond, Edmund is responding to the consideration that THIN ICE, we know more or less what we need to know to evaluate the rationality of his skating. That is because we know the practically relevant features: the consideration that THIN ICE is practically

¹¹⁷ See McDowell [2013: 19].

relevant to where one should skate on the pond. But being a fact is not a way of being practically relevant (or practically irrelevant). In general, if asked in what way the consideration that p bears on the question of whether A should Φ , it seems a kind of category mistake to respond with anything like, “In that it is true that p .” So while THIN ICE does alter our judgment of the rationality of Edmund’s action (*qua* consideration that Edmund acted upon), *its being the case* that THIN ICE should not.¹¹⁸

Rationality Without Good Reasons?

So I think there is a general problem for this first line of objection, whether we are thinking of rationality in course-grained or fine-grained terms, about what feature(s) of the situation one could point to as the right kind of evidence to overturn the intuitive judgment that Edmund’s action is rational. Since I do not think we will find any such evidence, I turn to a second line of objection to my main argument against the Factive View_N of epistemic reasons. This line of resistance concedes premise [2] of the argument, but tries to do so in a sense that does not imply that Edmund had good reasons to do what he did. For example, we may be able to pinpoint considerations or features that allow us to rationalize (in Davidson’s sense) Edmund’s action, or perhaps even pick out some rational requirement that the action satisfies. But these things do not strictly imply that there is any good reason for Edmund to be skating near the edge of the pond. So we can in this sort of way admit that Edmund’s action is in some sense rational while resisting premise [3] and rejecting the idea that Edmund acted for any *good* reason.¹¹⁹

¹¹⁸ Even if *being a fact* is not a way of being practically relevant, it might still be that *only facts* can have practical normative force. But I cannot see how that thought can be used here in a way that does not beg the question.

¹¹⁹ One similar idea that I do not address here is the intriguing suggestion that rationality is governed by requirements of coherence or consistency, and reasons-responsiveness is simply a separate phenomenon. I do not address this because I cannot see what coherence requirement a character like Edmund could be described as satisfying, such that his satisfaction of it leads to the judgment that he is rational. See Alex Worsnip [2015] for more on this idea.

I have only one thing to say in response to this sort of resistance, and it seems to me a hard point to maneuver around. It is, in many ways, the thought that makes the central argument of this chapter compelling (if indeed it is compelling). The thought is this: if we grant that Edmund is acting for the reason that THIN ICE, and that he is responding appropriately to that consideration, some subtlety is required to proceed to say that he has no good reason for doing what he is (or that there is no consideration that favors his action). If he indeed *has* good reason for doing what he is doing, it seems at least awkward to deny that *there is* good reason for doing what he is doing. Now it seems to me undeniable that in acting, Edmund is responding to the consideration that THIN ICE. That is more or less a stipulation of the case. It also seems to me undeniable that Edmund is responding appropriately to this consideration. The appropriate thing to do, in the light of the consideration that THIN ICE, is to avoid the middle of the pond when skating (if it helps, we may add that Edmund believes *that*, too). Indeed, as we saw earlier, it would be seem quite *inappropriate* to respond to this consideration by, say, skating all over the pond. Now, we can ask the question: what makes it appropriate to respond to the consideration that THIN ICE by skating near the edge of the pond? The answer that suggests itself is precisely that which would also suggest itself if it were the case that THIN ICE: the consideration that THIN ICE counts in favor of the response in question. But admitting that THIN ICE favors skating near the edge of the pond is tantamount to admitting that THIN ICE is a good reason for skating near the edge of the pond; and hence, in this case, tantamount to admitting that some good reasons are falsehoods.

It may be tempting here to make a psychologistic shift analogous to the one treated in the first three chapters: perhaps we should say that it is not THIN ICE that favors and makes appropriate Edmund's skating near the edge, but rather THIN ICE* (*Edmund believes* that the ice in the middle of the pond is thin). Such a shift would be ill-advised,

though. As I argued in the second chapter, every theory must allow a distinction between A 's Φ ing for the reason that r and A 's Φ ing for the reason that A believes that r . Once we countenance this distinction, it becomes clear that psychologizing the normative reason in this case will not help. Speaking on the psychologistic shift with respect to motivating reasons, Jonathan Dancy says that

This has the awkward consequence that in all such cases the agent makes a double mistake. In acting (as it seems to him) for the reason that p , he mistakenly takes it both to be the case that p and that he is acting for that reason; he is wrong on both counts. I, by contrast, remain unwilling to allow that in all cases where an agent makes a mistake, we need a different account of his reason for the relevant action. It still seems to me that the agent's reason is what it is, whether he is right about that or not.¹²⁰

I think that analogous points apply to normative reasons. For a clearer example, we can recall Dancy's illustration of the one who responds to the consideration that RATS: there are rats living in my shoes. The rational response to the consideration that RATS, it seems, is to call the pest control officer. But I have argued that this has a consequence some find unacceptable: it then seems difficult to deny that RATS *favors* calling the pest control officer, and therefore that some good reasons are falsehoods. Suppose instead that we say that what *really* favors calling the pest control officer is the consideration that RATS *: *I believe* that there are rats living in my shoes. If we then keep the case such that the agent calls the pest control officer for the reason that RATS, we accuse the agent of a double mistake: the agent is mistaken both in believing that RATS and in believing that RATS favors calling the pest control officer. This seems to commit us to saying that the agent makes a rational error in calling the pest control officer; but it seems clear that the agent makes no rational error in calling the pest control officer. On the other hand, if we admit that the rational response to the consideration that RATS is to call the pest control officer,

¹²⁰ See Dancy [2014: 89].

but say that the agent calls the pest control in the light of RATS *, we accuse the agent of a different double mistake: the agent is both mistaken about whether RATS, and also about what the consideration that RATS * favors (it surely favors calling the psychiatrist, rubbing one's eyes, getting some fresh air, or what have you).

There is surely a distinction between A 's Φ ing for the reason that r , and A 's Φ ing for the reason that A believes that r . The current line of resistance admits both that it is possible for A to Φ for the reason that r when it is not the case that r , and that in such cases A 's Φ ing might be rational. To resist the progression of my central argument in premise [3], we are looking for a way to allow all of this without admitting that in such cases, A Φ s for a good reason. It is simply not clear to me what coherent story can be told using the resources of a kind of Psychologism about normative reasons. In that case, the challenge raised earlier remains: provide and motivate accounts of *being the favored response* and *being the appropriate response* such that one can cleanly tease these things apart when dealing with characters like Edmund. I am skeptical that this can be done; and until it is done, I think that we should say that Edmund's skating near the edge of the pond is rational; and that is so because it favored by the false consideration that the ice in the middle of the pond is thin. Otherwise, we are left without a plausible account as to what makes his skating near the edge rational.

CAN THE FACTIVE TURN BE FORCED?

I think that the foregoing considerations land us back precisely where we were two sections ago: it is possible for an agent to act rationally in the light of the falsehood. If the agent acts rationally, there must be something in virtue of which that is so — there must be a rationality-making feature on the scene. Not only that, it seems, but there must be a rationality-making feature on the scene, which the agent responded to in acting. If we are

dealing with a case in which the agent acts only in the light of falsehoods, it seems that we are forced to choose our rationality-maker(s) from among these falsehoods. But then some rationality-making features, favoring features, or good reasons must be falsehoods.

There is one way of avoiding the conclusion of my argument that I have not discussed, and I should like to conclude the substance of this chapter by commenting on it. It is the piece of the first objection which I previously set aside for later. I know that there are many who have a sort of bedrock intuition that something that is not the case simply cannot favor anything. So long as we analyze good reasons (or normative reasons) as considerations that favor things, this amounts to a bedrock intuition that anything that is not the case simply cannot be a good reason. Some of those who have this thought may even find it to be more intuitive and plausible than each of the premises of my central argument here. Instead of resisting the cogency of the argument directly, then, one might claim that the conclusion of my argument serves up a *reductio*. The link that I worked to establish between practical motivating reasons and epistemic normative reasons could possibly be used against me in this way. I can see two general ways in which this might be tried. The first is that one might treat my argument as a kind of paradox: even though it cannot be said exactly *what* the flaw is in the argument, we know that *something* must be wrong with it given the conclusion it arrives at. The second possibility is that one might think that the implausibility of the conclusion of my argument should lead us to reject one of the premises in the argument — perhaps whichever seems to be the weakest link.

I am not sure how substantively I could proceed to debate my opponent who takes the first option here; but I also suspect that most will not be satisfied with that option. So I would like to close by making some remarks about the second option. I anticipate that most of my opponents will find the first premise to be the weak link here. My argument, in a sense, succeeds; but I have got it backwards. What the argument shows is not that the

Factive View_N of epistemic reasons is false, but rather that the Factive View_M of practical reasons is true. I would of course be pleased to receive the concession regarding the sense in which the argument succeeds. But this would be only a partial success if I intend to offer an argument for the Nonfactive View_N of epistemic reasons. If the crucial connection made in this chapter between the factivity of motivating reasons and normative reasons holds, the question then becomes which of the Nonfactive View_M and the Factive View_N is more plausible.

I myself feel the pull of the thought that something that is not the case cannot favor anything — though not as strongly as I feel the pull of the thought that one can act for a reason that is not the case. In any event, I argued at length for the Nonfactive View_M of practical reasons in the first chapter. What can be said in favor of the Factive View_N of epistemic reasons? One strategy is to subscribe to what might be called the Knowledge View of epistemic reasons, according to which: in order for r to be a reason for A to Φ , A must know that r . Giving due treatment to such a view would take us outside the bounds of this present project. At this moment, I should only like to make a small point about a case involving a Hornsby-McDowell style Edmund, who *truly* believes, for good reason, that the ice in the middle of the pond is thin. It seems to me deeply implausible to say in this case that Edmund's not *knowing* that the ice is thin precludes his skating near the edge of the pond from being rational. Perhaps this thought is ultimately correct, and some kind of Knowledge View of normative reasons is plausible. Still, though, it would not then be that that the Factive View_N, *per se*, is correct. On this sort of view, it would not be the *truth* of the consideration that THIN ICE which puts it in a position to favor Edmund's action. Rather, Edmund's *knowing* that THIN ICE is what allows it to favor his action. So we find here no resources for saying that the Factive View_N is correct in its own right.

I have begun to wonder just how independently plausible it is that something that is not the case cannot favor anything. Since at this point in the dialectic we are conceding that if the Nonfactive View_M is true, then the Nonfactive View_N is true, and I take myself to have made a strong case for the Nonfactive View_M, I should like to try to lay the foundation for understanding the favoring relation nonfactively. In particular, I shall offer a conception on which the favoring relation is doubly nonfactive: in order for a claim of the form “*P* favors *q*” to be true, it need not be the case that *p* or that *q*. In painting such a picture, I think an analogy to logical or probabilistic implication is useful. Consider claims like “*P* implies *q*” or “*P* entails *q*.” In order to make such claims — placing *p* and *q* in implication or entailment relations — we need not know whether *p* or *q* is indeed true. These relations hold independently of the truth-values of the items so related. I think that an appealing conception of the favoring relation works analogously: in order to make a claim of the form “ Φ favors Ψ ,” we need not know whether it is the case that Φ or Ψ . I find this idea to be especially appealing for normative reasons for belief. We might think that the consideration that THIN ICE favors believing that DANGER: it is dangerous to skate in the middle of the pond. For this to be so, need it be the case that THIN ICE? I think not. Indeed, one appealing thought might be that the very thing that makes rational to believe DANGER on the basis of the consideration that THIN ICE just is that THIN ICE implies or makes it more likely that DANGER. But as we have just observed, we need not say that it is the case that THIN ICE in order to place it in an implication or entailment relation to DANGER.

So, at least for normative reasons for belief, there is some initial appeal to the idea that claims of the form “*P* favors believing that *q*” are insensitive to the truth of *p*. Abstracting away from belief, can we not also say that, in general, claims of the form “*P* favors Φ ing” are insensitive to the truth of *p*? If not, why not — and what should we say

instead? There is clearly *some* relation of normative significance between, for example, THIN ICE and skating near the edge of the pond. Here is an alternative way of thinking, which perhaps threads together elements from all three lines of resistance consider so far. According to Jonathan Dancy,

To explain the action, we show how doing it can have made sense to the agent... what we show is that, relative to how the agent took things to be, it made good sense to do this action. If things had been as he supposed, this action would indeed have been the right or sensible thing to do.

There are two questions we can ask, then:

1. Did he do the right thing?
2. Had things been the way he supposed them to be, would what he have did been right?

The answer to the first question may be no even if the answer to the second is yes. Take a case where he (perhaps reasonably) believes that she would welcome his advances but is too shy to do much to encourage him. He acts in a way that would have been right if things had been as he supposed. But in fact things are not as he supposed, and what he does is not right (though perhaps pardonable, if his original belief was reasonable rather than, say, the result of lustful thinking).¹²¹

These remarks are made about what we might think of as *moral* reasons; but the same thoughts apply, Dancy holds, to other sorts of normative reasons:

To We see that the distinction drawn above for the moral case can be drawn more generally, in terms of what there is most reason to do. There are two questions:

1. Did he do what there was most reason to do?
2. Had things been the way he supposed them to be, would his action have been the one there was most reason to do?

We explain the action by showing that the answer to the second question is yes. We make sense of the agent's doing what he did, even when there was inadequate reason to do it, by showing that he would have had good reason had things been as he supposed.

¹²¹ Admittedly, there is some awkwardness about the second *relata* here. It is natural to think of the things that are favored as *responses* (bodily actions, attitudes, et cetera); and responses in the relevant sense cannot be true or false. We can set this aside, though, since our present concerns are about the first *relata*.

What this shows is that to explain an action is to justify it only in a certain sense. It is not to show that it was what there was most reason to do. It is to show that it would have been if the agent's beliefs had been true. But to show this much is to mount some defence of the agent for so acting. For the agent's behaviour is pardonable if (roughly) his beliefs were pardonable and had they been true he would have been doing what he ought, or what there was most reason to do.

This line of thought can perhaps countenance my analogy to logical implication, with a qualification. Talking about whether an action is justified is analogous not to talking about whether a proposition is entailed, but rather whether the entailed proposition is true. Even if p can entail q without being true, p cannot actually *make it true* that q unless p is true. Likewise, when we are talking of normative reasons, we are talking about features that *make it the case* that one should Φ , that the rational action is Φ ing, that the right thing to do is to Φ , or some other such. But nothing that is not the case can make anything else the case. Hence, normative reasons must be things that are the case. Now, what we have in cases like IGNORANT SKATING are considerations that *were they so, would* make it the case that one should Φ (or that it is rational to Φ , or some other such). This can explain why, even in situations where agents are wrong about the way the world is, we are tempted to describe their actions as rational. But in fact, this does nothing to show that these considerations *actually* favor anything when they are not so.

This is a compelling line of response, and I am almost satisfied by it. I will close by trying to say what I think is lacking in it. To the extent that I find this sort of picture unsatisfying when applied to characters like Edmund, the dissatisfaction has its roots in two intuitive judgments (already mentioned) that I find hard to shake off. The first is that seems hard to deny that Edmund *has* good reason for skating near the edge of the pond. But if *there is* no good reason for doing that, we should like to know how we can make sense of the idea that he had any good reason for doing it: how can one have something that there is not? It could be replied that what it is to have a good reason is to take

something to be the case, and for that thing to be such that *were* it the case, then it would indeed favor what the agent takes it to favor. But I find this thought not completely satisfying, since Edmund's action seems *actually* worthy of positive normative appraisal. It is not that *were* it the case that THIN ICE, then Edmund's skating near the edge *would have* been sensible. It is that, even things being such as they are, Edmund's action *is* sensible. So I think there is some pressure to say that whatever makes it the case that his action is sensible is also something such that *it actually does* make the action sensible — not that it would have made the action sensible if things were otherwise.

The second source of dissatisfaction, when applying Dancy's analysis to characters like Edmund, is that I find hard to shake off the intuitive judgment that Edmund's skating is not *irrational*. The analogue in the cases Dancy treats, I take it, would be that the action in question is *wrong* or *impermissible*, or that there was *inadequate reason* to do it (or perhaps *most reason* to do something else). In the cases Dancy is imagining, these verdicts are made more attractive by appeal to the thought that the actions, though strictly impermissible, might be *pardonable* or perhaps *excusable*. Though we can rationalize the actions by giving the reasons for which they were done, and in this way present a kind of defense of the actions even in cases where the motivating consideration is a falsehood, it remains that in these cases such explanations contain no real justification of the actions. In fact, we can see that there is justification for *not* doing what the agents in question did. So, though (for example) we could explain his advances by citing the motivating consideration that she would welcome his advances, if indeed she would not welcome his advances then we should say that what he did was impermissible (though perhaps pardonable).

The problem with applying this analysis to cases like that of Edmund's is that there are two important ways in which the cases are not analogous. The first is that, when

Edmund skates near the edge of the pond for the reason that the ice in the middle is thin, there is no hint that there is good reason for Edmund *not* to skate near the edge of the pond. In the examples Dancy treats, the wrong-making feature is conspicuous, since it is just the falsity or negation of the motivating consideration the agent takes to be the case. He is motivated by the consideration that she would welcome his advances; but in fact she would not welcome his advances, and that she would not welcome them is precisely what makes his action wrong. But this is not how things are in cases like IGNORANT SKATING. As we have already observed, it is not as though the fact that it is not the case that THIN ICE makes Edmund's skating near the edge impermissible. In fact, there does not seem to be any feature around that we could cite as that which makes Edmund's action irrational or impermissible. So not only does Edmund's action seem rationally permissible; there is also little to no suspicion that it might be impermissible. In the cases Dancy analyzes, we can be offered softening considerations for saying that what the agent did was impermissible, by saying things like, "Were things as the agent supposed, then it would have been the right thing to do. So the agent's action is perhaps pardonable or excusable." But this kind of softening consideration only has weight if there is some sense in which we can clearly establish that the agent should not have done what she did. It makes no sense to say, "What you did was right, but you may be excused for doing it." If there is no independent case to be made for the thought that Edmund should not have done what he did (or that there was no reason to do it), then softening considerations to the effect that his action is excusable should not warm us up to that thought. So it seems to me that we are left only with the judgment(s) that (for example) Edmund's skating near the edge of the pond is rational, that he has good reason to so skate, and that there is no feature of the case that counts against his so skating. In the light of these considerations,

what should prevent us from saying that there was good reason for Edmund to do what he did?

Summary Remarks

So, as things stand, I am inclined to follow my argument through to its conclusion — despite the admitted (though slight, in my view) discomfort in saying that r can favor A 's Φ ing even if it is not the case that r . I have made the case for the idea that A can Φ for the reason that r even if it is not the case that r . I have also made the case for the thought that when A so Φ s, A 's Φ ing can be rational, and so may be done for good reason. There may be ways of divorcing Φ ing rationally from Φ ing for good reason, but I cannot so far see a clean way of making the distinction that would provide good grounds for rejecting the idea that, for example, Edmund acts for a good reason when he skates near the edge of the pond. So, I think we should follow the natural way of thinking on which if A 's Φ ing is rational, then there must be some good reason for A to Φ — some consideration that counts in favor of A 's Φ ing. Since in cases like Edmund's, the only good candidate is the false consideration that motivates him to act as he does, it is plausible to say that some good reasons are falsehoods.

A PRACTICAL THEORY OF EPISTEMIC REASONS

Since I have endorsed the principle of UNIFORMITY, I think that all of this goes equally well if we speak only of reasons for belief. Indeed, I think we need not even rely on the principle to make the connection between cases like IGNORANT SKATING and cases of believing for reasons. For suppose that instead of evaluating the rationality of Edmund's skating near the edge of the pond, we evaluate the rationality of Edmund's belief that it is dangerous to skate in the middle of the pond. It seems even more appealing to me in this case to say that Edmund's response to the false consideration that

THIN ICE is rational. It is rational to believe, on the basis of the reasonably believed consideration that THIN ICE, that it is dangerous to skate in the middle of the pond. That it is not the case that THIN ICE makes no difference to this. It is tempting to think, though, that a belief can only be rational if it is held on the basis of those things that actually make it rational. Here is John Turri on the standard conception of the relation between what are called “doxastic” and “propositional” justification:

IF (i) p is propositionally justified for S in virtue of S’s having reason(s) R, and (ii) S believes p on the basis of R, THEN S’s belief that p is doxastically justified.¹²²

I take it is standard to think of the connection not merely as a conditional, but also as a sort of analysis of doxastic justification.¹²³ Using this taxonomy, it seems to me clear that we should say that Edmund’s belief that skating in the middle of the pond is dangerous is doxastically justified: he has reasons for believing it, and he believes it on the basis of those reasons.

So while I have used the assumption of UNIFORMITY to temporarily bridge the gap between so-called practical and epistemic reasons in this manuscript, I think that there is an independent case to be made for the Nonfactive View_N of reasons for belief.¹²⁴ In fact, I think that each of the main positions I have argued for with respect to reasons for action are independently just as plausible as positions about reasons for belief. So, I also think there is a good case to be made for the Nonfactive View_M applied to the reasons for which agents believe. Epistemologists tend not to talk of motivation; instead, discussion of the Nonfactive View_M may be hiding in the debate concerning the nature of

¹²² See Turri [2010: 314]. Turri proceeds to argue against this claim; but he acknowledges that it is almost universally held among epistemologists. See my footnote 22.

¹²³ For evidence of this, see the passages that Turri [2010: 313-314] cites, which include work from Jonathan Kvanvig, John Pollock, Joseph Cruz, Marshall Swain, Keith Korb, and Richard Feldman.

¹²⁴ This would put me at odds with those who think that only facts can be evidence. See Williamson [2000] and Littlejohn [2012] for two strong defenses of the factive position.

what is called the epistemic basing relation. Here is a standard conception of that relation, from Keith Allen Korcz:

The epistemic basing relation is the relation that holds between a reason and a belief if and only if the reason is a reason for which the belief is held.¹²⁵

So, perhaps defending the Nonfactive View_M in epistemology amounts to saying that the epistemic basing relation is not factive: it is not the case that, in order for *A*'s belief that *p* to be based on the consideration that *r*, it must be the case that *r*. I think that one could make a case for this position in much the same way that I have made the case for the Nonfactive View_M of reasons for acting here. This would have some interesting consequences worth exploring. To mention just one: one of the mainstream views about the epistemic basing relation is that it is a kind of causal relation. But if the basing relation is nonfactive — that is: if *A*'s belief that *p* can be based on the consideration that *r* even when it is not the case that *r* — then there is some serious pressure not to say that the basing relation is causal, since nothing that is not the case can cause anything.

There should also be a debate about reasons for belief analogous to the one I waded into in my second chapter here. An important question about reasons for belief is: in order for *A* to believe that *p* for the reason that *r*, what cognitive connection must *A* stand in to *r*? It seems to me that much too little has been said on this topic.¹²⁶ Now, the issue of whether reasons *are* beliefs has been given some treatment. According to John Turri, for example, “Epistemic reasons are mental states. They are not propositions or non-mental facts.” But according to David Enoch,

¹²⁵ See Korcz [2015: 1].

¹²⁶ See Enoch [2010], Littlejohn [2012], and Williamson [2000] for rare examples of those who have discussed the issue. Even they treat it only in passing, though, as a part of a larger project about the nature of *normative* reasons for belief.

When it comes to *your reasons*, or *what you take to be the epistemically relevant features of the circumstances*, it is quite clear... that your reason is that Adam is wrong, not that you believe that he is.¹²⁷

This debate, however, is clearly analogous to the debate I have marked out between Psychologism and Non-Psychologism. It is not analogous to the debate about the cognitive constraint on acting for a reason. I do think, though, that both the Belief View and Non-Psychologism are attractive positions about reasons for belief. In order for *A* to believe that *p* for the reason that *r*, *A* must believe that *r*; and when *A* so believes, *A*'s reason is (usually) *that r*, rather than that *A* believes that *r*.

In this manuscript I have defended a theory packaging together the Nonfactive View_M, the Belief View, Non-Psychologism, and the Nonfactive View_N of acting for reasons. I have tried to show that the best arguments from opponents should not persuade us away from this intuitive package of views; and in doing so I have given special attention to what I have called arguments from error. Such arguments should not compel us toward the Factive View_M, the Knowledge View, Psychologism, or the Factive View_N of acting for reasons. A complete theory of reasons and responding to reasons will not discuss only reasons for acting, but also reasons for believing (and perhaps other sorts of reasons as well). I have (merely) suggested here at the end that a plausible view on this second front would be one that packages the analogous Nonfactive View_M, the Belief View, Non-Psychologism, and Nonfactive View_N on reasons for belief. I also expect that objections analogous to those treated in this manuscript are no more promising in the realm of epistemic reasons. A complete theory of reasons would thoroughly lay out and defend this package of views for reasons for belief; I leave this more expansive project to future work.

¹²⁷ See Turri [2009: 490] and Enoch [2010: 983].

References

- Achinstein, Peter [1983]. *The Nature of Explanation*. Oxford University Press.
- Alvarez, Maria [2009]. "How Many Kinds of Reasons?" *Philosophical Explorations* 12 (2): 181-193.
- Alvarez, Maria [2010]. *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford University Press.
- Alvarez, Maria [2016]. "Reasons for Action: Justification, Motivation, Explanation." *Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). Summer Edition.
- Anscombe, G. E. M. [1957]. *Intention*. Harvard University Press. Second Edition, 1963.
- Anscombe, G. E. M. [1957b]. "Intention." *Proceedings of the Aristotelian Society* 57: 321-322.
- Anscombe, G. E. M. [1989]. "Von Wright on Practical Inference." In P. A. Schlipp and L. E. Hahn (eds.), *The Philosophy of George Henrik Von Wright*. Open Court: 377-404.
- Arpaly, Nomy [2003]. *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford University Press.
- Bach, Kent [1980]. "Actions are not Events." *Mind* 89 (353): 114-120.

Baier, Kurt [1958]. *The Moral Point of View*. Cornell University Press.

Bratman, Michael [1999]. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press.

Chisholm, Roderick [1970]. "Events and Propositions." *Nous* 44 (1): 15-24.

Collins, Arthur W. [1984]. "Action, Causality, and Teleological Explanation." *Midwest Studies in Philosophy* 9 (1): 345-369.

Comesaña, Juan and Matthew McGrath [2014]. "Having False Reasons." In Clayton Littlejohn and John Turri (eds.), *Epistemic Norms*. Oxford University Press: 59-80.

Cowper, Elizabeth A. [1992]. *A Concise Introduction to Syntactic Theory: The Government-Binding Approach*. University of Chicago Press.

Dancy, Jonathan [1995]. "Arguments from Illusion." *Philosophical Quarterly* 45 (181): 421-438.

Dancy, Jonathan [2000]. *Practical Reality*. Oxford University Press.

Dancy, Jonathan [2003]. "Replies." *Philosophy and Phenomenological Research* 67 (2): 468-490.

Dancy, Jonathan [2004]. *Ethics Without Principles*. Oxford University Press.

- Dancy, Jonathan [2008]. "On How to Act—Disjunctively." In Haddock, Adrian and Fiona Macpherson (eds.), *Disjunctivism: Perception, Action, Knowledge*. Oxford University Press 262-280.
- Dancy, Jonathan [2011]. "Acting in Ignorance." *Frontiers of Philosophy in China* 6 (3): 345-357.
- Dancy, Jonathan [2014]. "On Knowing One's Reason." In *Epistemic Norms*. Clayton Littlejohn and John Turri (eds.). Oxford University Press, 2014: 81-96.
- Darwall, Stephen [1983]. *Impartial Reason*. Cornell University Press.
- Darwall, Stephen [2003]. "Desires, Reasons and Causes." *Philosophy and Phenomenological Research* 67: 435-443.
- Davidson, Donald [1963]. "Actions, Reasons, and Causes". *Journal of Philosophy* 60 (23) 685-700.
- Davidson, Donald [1980]. *Essays on Actions and Events*. Oxford University Press. Second Edition, 2001.
- Davis, Wayne A. [2003]. "Psychologism and Humeanism." *Philosophy and Phenomenological Research* 67 (2): 452-459.

Enoch, David [2010]. "Not Just a Truthometer: Taking Oneself Seriously (but not Too Seriously) in Cases of Peer Disagreement." *Mind* 119 (476): 953-997.

Gibbard, Alan [1990]. *Wise Choices, Apt Feelings*. Harvard University Press.

Haegeman, Liliane [1991]. *Introduction to Government and Binding Theory*. Blackwell.

Hawthorne, John and Jason Stanley [2008]. "Knowledge and Action." *The Journal of Philosophy* 105 (10): 571-590.

Hempel, Carl G. and Oppenheim, Paul [1948]. "Studies in the Logic of Scientific Explanation." *Philosophy of Science* 15 (2): 135-175.

Hieronymi, Pamela [2011]. "Reasons for Action." *Proceedings of the Aristotelian Society* 111 (3): 407-427.

Hornsby, Jennifer [2004]. "Agency and Actions." *Royal Institute of Philosophy Supplement* 55: 1-23.

Hornsby, Jennifer [2007]. "Knowledge, Belief, and Reasons for Acting." In C. Penco, M. Beaney, and M. Vignolo (eds.), *Explaining the Mental*. Cambridge Scholars Publishing, 88-105.

- Hornsby, Jennifer [2008]. "A Disjunctive Conception of Acting for Reasons." In *Disjunctivism: Perception, Action, Knowledge*. Adrian Haddock and Fiona Macpherson (eds.). Oxford University Press, 2008: 244-261.
- Hyman, John [1999]. "How Knowledge Works." *The Philosophical Quarterly* 49 (197): 433-451.
- Hyman, John [2011]. "Acting for Reasons: Reply to Dancy." *Frontiers of Philosophy in China* 6 (3): 358-368.
- Hyman, John [2015]. *Action, Knowledge, and Will*. Oxford University Press.
- Kelly, Thomas [2008]. "Evidence: Fundamental Concepts and the Phenomenal Conception." *Philosophy Compass* 3 (5): 933-955.
- Kelly, Thomas [2014]. "Evidence." *Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). Summer Edition.
- Korcz, Keith Allen [2015]. "The Epistemic Basing Relation" *Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). Summer Edition.
- Lenman, James [2011]. "Reasons for Action: Justification vs. Explanation." *Stanford Encyclopedia of Philosophy*. Winter Edition.

Littlejohn, Clayton [2012]. *Justification and the Truth-Connection*. Cambridge University Press.

Littlejohn, Clayton [2013a]. "No Evidence is False." *Acta Analytica* 28 (2): 145-159.

Littlejohn, Clayton [2013b]. "The Russellian Retreat." *Proceedings of the Aristotelian Society* 113 (3): 293-320.

Littlejohn, Clayton [2017]. "A Plea For Epistemic Excuses." Forthcoming in F. Dorsch and J. Dutant (eds.), *The New Evil Demon*. Oxford University press.

Mantel, Susanne [2014]. "No Reason for Identity: on the Relation between Motivating and Normative Reasons." *Philosophical Explorations* 17 (1): 49-62.

McDowell, John [2013]. "Acting in the Light of a Fact," in *Thinking About Reasons: Themes from the Philosophy of Jonathan Dancy*, David Bakhurst, Brad Hooker, and Margaret Olivia Little (eds.). Oxford University Press: 13-28.

Mele, Alfred [2007]. "Reasonology and False Beliefs." *Philosophical Papers* 36 (1): 91-118.

Mele, Alfred [2013]. "Actions, Explanations, and Causes"" in G. D'Oro and C. Sandis (eds.), *Reasons and Causes: Causalism and Anti-Causalism in the Philosophy of Action*. Palgrave Macmillan.

- Moore, G. E. [1903]. *Principia Ethica*. Second Edition, T. Baldwin (ed.). Cambridge University Press, 1993.
- Muñoz, Daniel [2017]. "Thinking, Acting, Considering." *Australasian Journal of Philosophy*.
Forthcoming.
- Parfit, Derek and John Broome [1997]. "Reasons and Motivation." *Proceedings of the Aristotelian Society* 71: 99-146.
- Parfit, Derek [2001]. "Rationality and Reasons." In Dan Egonsson, et al (eds.), *Exploring Practical Reasons*. Ashgate Press: 17-39.
- Parfit, Derek [2011]. *On What Matters*. Oxford University Press.
- Prichard, H. A. [2002]. "Duty and Ignorance of Fact." In his *Moral Writings*, J. MacAdam (ed.). Clarendon Press.
- Pryor, James [2007]. "Reasons and That-Clauses." *Philosophical Issues* 17 (1): 217-244.
- Putnam, Hilary [1973]. "Meaning and Reference." *Journal of Philosophy* 70 (19): 699-711.
- Raz, Joseph [1975]. *Practical Reason and Norms*. Hutchinson.
- Raz, Joseph [1999]. *Engaging Reason: On the Theory of Value and Action*. Oxford University Press.

Raz, Joseph [2011]. *From Normativity to Responsibility*. Oxford University Press.

Ross, W. D. [1930]. *The Right and the Good*. Clarendon Press.

Ryle, Gilbert [1949]. *The Concept of Mind*. 60th Anniversary Edition. Routledge, 2009.

Sandis, Consantine [2013]. "Can Action Explanations Ever Be Non-Factive?" in *Thinking About Reasons: Themes from the Philosophy of Jonathan Dancy*, David Bakhurst, Brad Hooker, and Margaret Olivia Little (eds.). Oxford University Press: 29-29.

Scanlon, T. M. [1998]. *What We Owe to Each Other*. Harvard University Press.

Schueler, G. F. [2003]. *Reasons and Purposes: Human Rationality and the Teleological Explanation of Action*. Oxford University Press.

Schnee, Ian [2015]. "The Explanatory Argument for Factualism." Unpublished Manuscript.

Schroeder, Mark [2008a]. "Having reasons" *Philosophical Studies* 139 (1): 57-71.

Schroeder, Mark [2008b]. *Slaves of the Passions*. Oxford University Press.

Sehon, Scott [2005]. *Teleological Realism: Mind, Agency, and Explanation*. MIT Press.

Setiya, Kieran [2007]. *Reasons Without Rationalism*. Princeton University Press.

Smith, Michael [1992]. "Valuing: Desiring or Believing?" In Charles, David and Kathleen Lennon (eds.), *Reduction, Explanation, and Realism*. Oxford University Press, 323-360.

Smith, Michael [1994]. *The Moral Problem*. Blackwell.

Smith, Michael [1998]. "The Possibility of Philosophy of Action", in J. Bransen and S. Cuyper (eds.), *Human Action, Deliberation and Causation*. Kluwer Academic Publishers.

Smith, Michael [2010]. "Humeanism About Motivation", in T. O'Connor and C. Sandis (eds.), *A Companion to the Philosophy of Action*. Wiley-Blackwell.

Strawson, P. F. [1992]. *Analysis and Metaphysics: An Introduction to Philosophy*. Oxford University Press.

Turri, John [2009]. "The Ontology of Epistemic Reasons." *Nous* 43 (3): 490-512.

Turri, John [2010]. "On the Relationship between Propositional and Doxastic Justification." *Philosophy and Phenomenological Research* 80 (2): 312-326.

Unger, Peter [1975]. *Ignorance: A Case for Scepticism*. Oxford University Press, 2002 Edition.

Williams, Bernard [1981]. *Moral Luck*. Cambridge University Press.

Williamson, Timothy [2000]. *Knowledge and its Limits*. Oxford University Press.

Williamson, Timothy [201X]. “Acting on Knowledge.” Forthcoming in J.A. Carter, E. Gordon, and B. Jarvis (eds.), *Knowledge-First*. Oxford University Press.

Worsnip, Alex [2015]. “The Conflict of Evidence and Coherence. *Philosophy and Phenomenological Research*.”