

# ACTING FOR REASONS

by

**Laura Tomlinson**

B.A., Reed College, 2011

Submitted to the Graduate Faculty of  
the Kenneth P. Dietrich School of Arts & Sciences in partial  
fulfillment

of the requirements for the degree of

**Doctor of Philosophy**

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH  
DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Laura Tomlinson

It was defended on

July 29th 2019

and approved by

Michael Thompson, Professor of Philosophy

John McDowell, Distinguished University Professor of Philosophy

Jennifer Whiting, Distinguished Professor of Philosophy

Karl Schafer, Professor of Philosophy, University of California, Irvine

Dissertation Director: Michael Thompson, Professor of Philosophy

Copyright © by Laura Tomlinson  
2019

## ACTING FOR REASONS

Laura Tomlinson, PhD

University of Pittsburgh, 2019

This dissertation addresses the question, “What is the relation of reasons to actions?” Throughout this work I argue that reasons are internal to actions; that there is no specifying an action without reference to reasons; and thus that there is no thought of action apart from thought of reasons.

I begin in Chapter One by sketching a dialectic in the literature on the theoretical inference, showing that the philosophical understanding of it is beset by a serious dilemma between over-intellectualization, and under-intellectualization. Turning to the topic of reasons for acting, I argue that the relation between reasons, as understood by contemporary authors, and actions is beset by an analogous dilemma.

In Chapter Two, I examine two recently proposed solutions to the dilemma for the theoretical inference, with an eye toward what these solutions, though distinct, have in common. I argue that they share a similar kind of re-envisioning of the question of inference. Then I show that an analogous re-envisioning is not only possible in the practical case, but is already underwritten and embraced by a particular theory of action: the Anscombean theory. If that theory is in a privileged position to solve the dilemma that faces accounts of reasons for action, then we have good reason to accept it.

I turn to an ethical application of this conception of action in Chapter Three. Taking up the topic of moral worth—as opposed to moral desirability—I argue that a parsimonious and fully accurate account can be given only if we make use of the insights gained by the Anscombean theory of action. That theory is, therefore, bolstered still further by solving the puzzle of moral worth, as well as puzzles about *de dicto* and *de re* motivation.

## TABLE OF CONTENTS

<b>PREFACE</b> . . . . .	vii
<b>I. INTRODUCTION</b> . . . . .	1
<b>II. ACTING FOR REASONS ON THE MODEL OF THE THEORETI- CAL INFERENCE</b> . . . . .	6
A. Theoretical Inference: a prelude to the question of acting for a reason . . . . .	6
B. Transition to the proper topic . . . . .	12
C. Reasons for action and acting for a reason . . . . .	13
1. Reasons Fundamentalism . . . . .	15
a. Scanlon’s account of acting for a reason . . . . .	16
b. Raz’s account of acting for a reason . . . . .	19
2. The Humean Theory of Reasons . . . . .	22
D. On the value of the analogy . . . . .	25
E. Conclusion . . . . .	28
<b>III. A NON-STANDARD APPROACH TO INFERENCE AND ACTING FOR REASONS</b> . . . . .	29
A. Introduction . . . . .	29
B. The Non-Standard Accounts of Inferring . . . . .	32
1. Hlobil’s Force Account . . . . .	32
2. Valaris’s Semantic Account . . . . .	34
3. Lessons Learnt . . . . .	36
a. From causal process to act . . . . .	36
b. From atomism to complexity . . . . .	39

c.	An act that contains its own grounds . . . . .	40
C.	A Non-Standard Story of Action . . . . .	42
1.	The Standard Story of Action . . . . .	43
2.	The Non-Standard Story in Three Movements . . . . .	44
a.	First movement: rejection of a necessary causal relation be- tween reason and action . . . . .	44
b.	Second movement: focus on the internal complexity of actions . . . . .	47
c.	Third movement: actions contain their own grounds . . . . .	49
D.	Prospects of the Non-Standard Accounts . . . . .	53
E.	Morals of the (Non-Standard) Story . . . . .	58
F.	Conclusion . . . . .	60
<b>IV.</b>	<b>THE MORAL WORTH OF INTENTIONAL ACTIONS . . . . .</b>	<b>61</b>
A.	Introduction . . . . .	61
B.	Three Right Reasons Theories . . . . .	63
1.	Nomy Arpaly on Moral Worth . . . . .	63
2.	Julia Markovits on Moral Worth . . . . .	65
3.	Amy Massoud on Moral Worth . . . . .	66
C.	Acting Well as a Doctrine of Moral Worth . . . . .	69
D.	Problem Cases . . . . .	75
E.	Motivation <i>De Dicto</i> and Motivation <i>De Re</i> . . . . .	76
F.	Conclusion: Non-Accidentality . . . . .	83
<b>V.</b>	<b>CONCLUSION . . . . .</b>	<b>87</b>
	<b>BIBLIOGRAPHY . . . . .</b>	<b>95</b>

## PREFACE

To John McDowell, an ever-trusted leader; to Jennifer Whiting, a model and support; to Jonathan Dancy, a spirited objector; and to Karl Schafer, a knight in shining armor, I am deeply grateful. To Michael Thompson, I owe more than I can say—his work has been a continual inspiration, and his time and care have made me the philosopher I am today. I cannot thank him enough for everything he has given me.

For philosophical aid and stimulation, I am grateful to the many classmates I've had at Pitt; to Jonas Held, a couple of conversations with whom sparked many of the ideas in this dissertation; and to Milo Philips-Brown, an indispensable interlocutor. To the friends I made in Pittsburgh, a group of people whose virtues are boundless, I owe my sanity throughout these long years: to Justin Stewart, Charles Bettendorf, Tyke Nunez, Laura Davis, Preston Stovall, David DeBruijn, and Mike Strati. For unwavering support, as ever, I am grateful to my family. Without Julia, in particular, I don't know what would have become of me.

Most of all, however, I am grateful to and for Steve Makin, who has made me not just a better philosopher, but a better person as well. If I have any good ideas, I owe them to him.

## I. INTRODUCTION

When philosophers write about “practical reasoning,” there are two different things they might mean. On the one hand, there is what we might also call “practical deliberation:” the conscious thought-process whereby an agent arrives at an answer to the question, “what to do?” This sort of practical reasoning is familiar from daily life; it is often presented in novels, sometimes in voice-overs in films; if the question of what to do is hard enough, it is the kind of thing one might share with a parent or other trusted advisor.

This sort of practical reasoning is well worth whatever philosophical attention might be given to it. More often than not, it is only passingly treated, or contemplated by accident: its familiarity gives it a kind of magnetic pull on the minds of philosophers writing on other topics. But it holds up to sustained philosophical scrutiny as well: Jonathan Dancy, for instance, has given an illuminating taxonomy of the kinds of relevance reasons might have, showing that, if we attend carefully to concrete instances of practical reasoning, we can recognize a diversity of roles that reasons or facts may play in our deliberation.<sup>1</sup> Understanding better how we deliberate promises, among other things, to elucidate our conception of the values that we deliberate about.

But there is something else we might mean by “practical reasoning,” something that inhabits a more rarefied philosophical air. By “practical reasoning” philosophers sometimes mean to indicate the way in which practical things—actions, first and foremost—are reasoned, or related to reason, or expressions of a capacity for reason. Here, the focus is not on the process of reasoning that leads to action, but rather on what we might think of as the intersection point between reason and the practical realm; the question is not about what happens when we reason our way toward an action, or when we act on the basis of reasoning,

---

<sup>1</sup>See [Dancy \[2018\]](#).



but rather about how the concepts of *action* and *reason* meet. Perhaps this topic would be better called “practical reason:” reason in a practical key; reason put into a practical gear.

It is this idea, practical reason, that is the topic of this dissertation. My goal in the essays that follow is to examine, from three different angles, what it means for reason to be made practical.

Of course, just how we put this question controls much of what might be offered as an answer. At its heart, the question is a version of that most familiar of philosophical questions: what is the place of mind in the world? The question of acting for reasons—or, as I prefer to think of it, the question of action *simpliciter*—can be seen as a particularly vivid crisis of this question. Actions are paradigmatically worldly things: by their nature, they take place in the world, that is, in the causal order; indeed they depend on that causal order for their intelligibility. And yet, they are also, by their nature, the products of human reason; and they partake of the special normativity that only human things do.

There are many more determinate ways of asking this question about action—the problem of free will, in particular, comes to mind. But I have been more impressed by the ways in which the question of acting for a reason can be made to seem continuous with questions of mind and world as they arise in, as it is sometimes put, theoretical philosophy. If acting is interesting here insofar as it is an expression of our capacity to reason, then it seems plausible that headway should be able to be made by considering its similarity to broader or, perhaps, more familiar questions about that capacity.

When I first took up this topic, I was accustomed to thinking of it in terms of the *practical inference*. After all, if the puzzle concerns practical *reason*, it seemed suitable to grasp it in terms of a more determinate expression: namely, inference. Furthermore, the practical inference—introduced by Aristotle but notably discussed by Elizabeth Anscombe—itself seemed to me to isolate the idea of the formal quality of something that is both practical and the output of reason, rather than attracting thoughts of deliberative processes and the like. And, conceiving of the question as about the practical inference, I thought that philosophical treatment of the much more familiar theoretical inference would be the proper place to look for the beginnings of an account.

Little did I know that the theoretical inference is itself the object of much consternation

among a very active set of philosophers today. And yet the more I came to understand the contemporary debate about the theoretical inference, the more convinced I became of the identity of the topics. Theorists are accustomed to thinking of the practical inference as beset by its own special problems—oftentimes concluding that the idea of the practical inference is, in fact, hopeless compared to its theoretical counterpart. But it seems to me that the theoretical inference is itself on rather more unstable ground than such theorists have supposed. Indeed, it seems to me that once we recognize the issues afflicting the theoretical inference, we can see that both kinds of inference are, in fact, problematic in just the same ways. The problem of the practical inference—or practical reason, or practical reasoning—then opens up onto a much broader problematic: that of the idea of a rational process *as such*. Or at any rate, so I argue.

In the first essay of this dissertation, I present the analogy between the theoretical and practical cases as concerns the problem of inference. Contemporary theorists are not, on the whole, interested in the idea of the practical inference (concluding, as I said, that it is hopeless!); but they are very interested in the idea of reasons for acting. But theoretical inference can be understood to be one way that we believe for reasons; so it seems that practical reason can also be suitably grasped through the idea of acting for reasons. I argue that the problems that arise for the theoretical inference arise also for two currently dominant theories of reasons: the reasons fundamentalism espoused by T. M. Scanlon, and the Humeanism about reasons made popular by Michael Smith. Both of these conceptions of reasons, I argue, fall prey to objections analogous to those leveled at the dominant conceptions of theoretical inference. The upshot is two-fold: first, if my arguments against the two conceptions of reasons are sound, that shows that both of them would be unable to explain how we could ever act for reasons in their sense. This would, I think, be a fatal objection to a theory of reasons—even without accepting anything like internalism about reasons.<sup>2</sup>

The second upshot is that the problem of acting for a reason is not a special problem in need of special psychological or physiological explanation, but is, rather, the same sort of problem as the problem of theoretical inference. In other words, action, as much as belief,

---

<sup>2</sup>That is, we need not claim that in order for something to be a reason, one must be able to act for it; we need only think that sometimes people act for their reasons, and thus that an account of reasons must be able to accommodate that fact.

presents the puzzle of being something that is both *real*—that is, a part of human life—and rational—connected, roughly speaking, to the laws of reason.

But that essay is largely negative in its aim: to show that the idea of acting for reasons is under serious threat. In the second essay, my aim is positive. I argue there that a certain conception of intentional action—the non-standard story developed by theorists under the influence of Elizabeth Anscombe—holds special promise in addressing the problems laid out in the first essay. I begin by surveying recent accounts of theoretical inference that are aimed at solving the problems afflicting the theoretical inference. Though these accounts are importantly distinct, I think that what they share in common is the more striking. Both accounts, as I bring out, represent a shift away from understanding inference as a causal process that leads from some belief-states to some subsequent belief-state, and toward understanding inference as a special, complex kind of act that unifies conclusion with grounds.

I argue that the Anscombean account of intentional action already has precisely this form: it understands action not merely as the effect of a certain causal process, but as an internally complex whole that is constituted by its grounds. If theorists investigating the nature of theoretical inference are looking to solve the problems that arise for the theoretical inference by adopting an account with this form; and if the idea of acting for reasons is open to analogous problems; then the fact that one theory of action—the Anscombean theory—already has that form in place for conceiving of intentional actions counts, I claim, as good reason to accept that theory. In other words, I aim to show that the Anscombean theory of action is our best hope of understanding the relation of reasons to actions.

Illuminating, as it does, the relation of reasons to actions, the Anscombean theory is also, I claim, in a position to offer solutions to ethical problems that arise concerning reasons and actions. In the third essay, I make a study of this. I take up the problem of moral worth, which has traditionally been understood to be the moral property of actions determined precisely by the quality of the reasons for which they were performed. Hence, the thought goes, two agents may perform the same action, A, but for different reasons; and though A will be morally good to some fixed extent, the two performances might have different values of moral worth. But accounts of moral worth in terms of reasons tend to go awry in explaining the many intuitions we have about the moral worth of various actions. I argue that

one insight of the Anscombean view—that reasons for an action correspond to descriptions under which the action is intentional—can, by itself, provide a satisfactory account of moral worth.

Adopting a slightly-un-Anscombean form of expression, I argue that moral worth can be understood to be the moral goodness of a token action, which instantiates multiple types; whereas moral goodness, on the traditional conception, applies only to an action insofar as it instantiates a single type. In other words, once we bring into focus the fact that actions are *themselves* intentional under multiple descriptions—that is to say, that actions are constituted by strings of rationalized action-forms—we can appreciate how the materials provided by the action itself can explain the its moral worth. We do not, then, have to advert to reasons as separate entities, accidentally related to the action in question, in our account.

All in all, then, this dissertation offers a defense of an Anscombean theory of action—in particular, a conception of action that understands actions as constituted by their reasons and, therefore, inseparable from those reasons. Understanding the concept of action in this way fits it most comfortably—most elegantly, even—into the nexus of concepts in which it belongs.

## II. ACTING FOR REASONS ON THE MODEL OF THE THEORETICAL INFERENCE

### A. THEORETICAL INFERENCE: A PRELUDE TO THE QUESTION OF ACTING FOR A REASON

There has lately arisen a renewed interest among philosophers in the question of inference—the question being, what is it?<sup>1</sup> Inference, loosely, is a process whereby a thinker moves from belief in some proposition or propositions to belief in another.<sup>2</sup> The interest of this particular process to philosophers is its connection to rationality: a belief formed on the basis of inference is justified, and its justification is known to the thinker. Indeed, it can come to seem that inferring must be *the* central mechanism of rationality: any belief that is rationally held must be susceptible to a certain kind of scrutiny—we must be able to ask *why* it is held, and receive an answer which gives a reason to hold it. Inferring takes one from apprehension of a reason to believe something to having that belief. The question of inference, then, asks after the nature of a specially rational process that makes, so to speak, a difference in the life of a thinker. What the philosopher wants to know is: how does this work?

---

<sup>1</sup>See, for example, [Boghossian \[2014\]](#), [Neta \[2013\]](#), [Broome \[2012\]](#), [Hlobil \[2014\]](#), [Wright \[2012\]](#).

<sup>2</sup>There does not, unfortunately, appear to be much linguistic consensus on this issue and the surrounding issues. I am taking ‘inference’ in the way specified and examined below; this is to be distinguished from what I would call ‘deduction’, or the identification of logical relations between propositions, unrelated to their instantiation in the beliefs of any thinker; and it is to be distinguished from what we could call, more carefully, ‘reasoning.’ *Reasoning*, in this sense, would amount to a belief-involving process, but would accommodate, for example, the thinker considering a set of premises and the conclusion that follows, and rejecting one of the premises as a result. *Inference*, in my sense, is more basic than that: it is, as it were, the most basic process of rational belief-formation. In that bit of ‘reasoning’ I just described, an inference would take place in which the thinker formed the belief that “if one accepts those premises, one will be forced to accept that conclusion”—that is, an inference with hypothetical contents.

Paul Boghossian, whose article “What is Inference?” sets the stage for contemporary debate on the topic, organizes his discussion around a formulation of Frege’s: “To make a judgment because we are cognisant of other truths as providing a justification for it is known as *inferring*.”<sup>3</sup> Making minor modifications to Frege’s characterization, Boghossian defines inferring as follows: “S’s inferring from p to q is for S to judge q *because* S *takes* the (presumed) truth of p to provide support for q.”<sup>4</sup> This definition in turn imposes a condition on any account of inference, which Boghossian calls the *Taking Condition*: “Inferring necessarily involves the thinker *taking* his premises to support his conclusion and drawing his conclusion *because* of that fact.”<sup>5</sup>

The intuitive appeal of this conjunctive condition is clear enough. On the one hand, causal relations between beliefs can hold in the absence of apprehension of a justificatory relation between those beliefs—and in such cases, no inference, or rational transition, will have been made. Consider a thinker who judges that everyone at the party is having a great time, which makes him form the subsequent belief that nobody likes him (perhaps because of a depressive tendency of thought). He does not *take* his ‘premise’—that everyone is having a good time—to support his ‘conclusion’—that nobody likes him; he knows there is no rational connection between the two propositions, but nonetheless he comes to believe the latter because of his belief in the former. Though his belief is caused by another belief of his, it is not formed *on the basis* of that belief—it reflects a mere disposition of thought. Surely, then, he does not infer.

On the other hand, a thinker can apprehend a relation of support between two of his beliefs without having formed the supported belief on the basis of the other—and here too I think we will say he has not inferred. Suppose an agent believes that certain advances in technology indicate that humans will soon live on Mars; yet he believes that humans will soon live on Mars because he read it in a book as a child (“In the year 2020, man finds a new home among the stars...”), and so came to believe it long before those technological advances came about. Here the agent believes that his premise—that certain technological advances have been made—supports his conclusion—that humans will soon live on Mars; but he does

---

<sup>3</sup>quoted on 4 in [Boghossian \[2014\]](#)

<sup>4</sup>*ibid.*, 4

<sup>5</sup>*ibid.*, 5

not believe the conclusion *because* of the premise—after all, he believed it already. So again, this thinker does not infer.

So, Boghossian concludes, any account of inferring is subject to the Taking Condition: the thinker must *take* his premises to support his conclusion and draw his conclusion *because* of that fact. But to attempt to give an account of inference by working out just what is required by the Taking Condition proves problematic.

First of all, we need to understand just what it means to “take” something to be the case. The most obvious candidate would be that to *take* a proposition to be the case is, simply, to *believe* it. The Taking Condition would then require that a thinker *believe* that his premises support his conclusion and that he draw his conclusion because of this belief.

But objections to this construction immediately arise. The first is a kind of Sophistication Objection.<sup>6</sup> It seems that there are thinkers to whom we would want to attribute the capacity for inference but for whom forming a belief with the requisite content would be too sophisticated. Children, for example, seem perfectly capable of reasoning, and so drawing inferences. But believing that some proposition supports another involves having and employing the concept of *support*—in particular, *rational support*—and it seems implausible to maintain that children possess such a concept. Thus, inferring cannot require forming a belief that one’s premises support one’s conclusion.<sup>7</sup>

But perhaps we can set aside what might be thought to be fringe cases, about whom Sophistication worries arise—after all, such thinkers prove problematic for a wide variety of philosophical theories, and so our conception of inference need not be especially beholden to them.<sup>8</sup> Still, however, we must give an account of the *because* in the Taking Condition. How, in other words, would a taking of this kind, understood as some sort of representational state concerning justification, effect the drawing of the conclusion?

One possibility—given that it seems we must rely on the belief about justification just as we must rely on the premises of an inference—is that we register the taking-belief simply *as* another premise. But putting the belief about support on a par with the premises as

---

<sup>6</sup>See [McHugh and Way \[2016\]](#)

<sup>7</sup>Boghossian gives this objection himself: pp 6-7; see also [McHugh and Way \[2016\]](#), 6.

<sup>8</sup>Though children are probably not the only ones—as Barbara Winters puts it, most normal people are probably “epistemically unaware” ([Winters \[1983\]](#), 217).

just another premise leads, famously, to a vicious regress. Lewis Carroll’s Tortoise made this point more than a century before Boghossian proposed the Taking Condition. The objection runs as follows: suppose we are attempting to specify an inference from premises A and B to conclusion Z. Now formulate the proposed taking-belief relating A and B together with Z as C: “if A and B be true, Z must be true.”<sup>9</sup> Inferring Z, then, relies on the premises A, B, and C, since it is supposed to rely not only on the premises, but also on the cognizance of the premises’ justification of the conclusion, represented by belief in C. But now it seems that in order to understand this subject as inferring Z from A, B, and C, she must be cognizant of A, B, and C jointly justifying Z—in other words, we must ascribe to her yet another belief, call it D, that states that if A, B, and C are true, then Z must be true. And off the regress goes.

This leaves the import of the Taking Condition an open question. At any rate it seems clear that the Taking Condition cannot, itself, *give* an account of inference, since its articulation leads to the difficulties just reviewed. Boghossian himself concludes his discussion of inference with the admittedly inconclusive suggestion that rule-following, understood as an unanalyzable primitive, might provide the right account. But authors before Boghossian—as well as in his wake—have put forward a different response to these issues, having already been faced with Carroll’s regress. The dominant alternative to accounts that emphasize the Taking Condition, like Boghossian’s, is Dispositionalism about inference. Dispositional accounts take the lesson of Carroll’s parable to be that reasoning must in some sense be blind—precisely the denial of the Taking Condition. Setting *taking* aside, such accounts then aim to understand the nature of the *causal* relation between acceptance of premises and acceptance of a conclusion. Inferring, then, is to be understood first and foremost as a causal process in which one mental state (or act) causes another.<sup>10</sup> In order to allay worries about deviant causal chains, authors posit mental dispositions—dispositions of a thinker’s internal mental life—as the causal mechanism at stake.<sup>11</sup> Barbara Winters provides an early articulation of Dispositionalism: “A infers  $q$  from other beliefs  $p$  only if the set of beliefs

---

<sup>9</sup>Carroll [1895], 432

<sup>10</sup>Wedgwood [2006], for example, declares this at the outset of his investigation of the inference.

<sup>11</sup>On deviant causal chains—in particular, external ones, see Wedgwood [2006]. It is worth noting, in this regard, that ‘dispositions’ here are understood as a kind of abstract, widely applicable, form of efficient causation. The important aspect, for our purposes, is that dispositions are supposed to be non-normative.



[ $p, q$ ] instantiates an inference pattern and  $A$ 's transition from  $p$  to  $q$  is a result of  $A$ 's general disposition to make transitions that exhibit that form.”<sup>12</sup>

This formulation, however, is immediately unsatisfying. As it stands, the fact that the antecedent beliefs provide justification for the conclusion is wholly external to the disposition manifested by the agent. It is, in other words, as good as accidental that the thinker's disposition instantiates a valid inference pattern at all. But if inference is to be understood as the manifestation of a disposition, surely it should be distinguishable from other dispositions of thought in a more robust manner than this. For a transition to amount to an inference, it is not good enough that the antecedent beliefs both cause the conclusion beliefs *and* rationalize them; we feel they must cause them *in virtue of* rationalizing them.

Ralph Wedgwood has more recently addressed this problem of “causation in virtue of rationalization”—which slogan represents his solution to the problem as well. As he notes, we can make sense of one event's causing another *in virtue of* some property of the first in typical, physical causal contexts, so it should not be mysterious here. So, for example, running a marathon might have caused one to collapse not in virtue of muscle fatigue, but in virtue of dehydration. So too in the mental case, then, we can say that one mental state causes another in virtue of rationalizing the latter. Hence, the disposition one manifests in drawing an inference “must be one that can be specified by means of a function that maps the stimulus event-type *coming to be in some mental states or other that rationalize forming a belief in  $p$*  onto the response event-type *forming a belief in  $p$* .”<sup>13</sup> Such a disposition, which responds to “*rationalizers as such*,” Wedgwood deems to be an “*essentially rational disposition*.”<sup>14</sup>

Now, Wedgwood's description of this disposition seems ambiguous between two possibilities—for it is not clear whether the *rationalizing* captured in the phrase “that rationalize a belief in  $p$ ” is supposed simply to be a fact about the stimulus mental states, or whether it is supposed to be contained in their content. On the former reading, the disposition would take as an input some belief in  $p$ , simpliciter, and result in belief in  $q$ ; and it would be—independently, as it were—a fact that belief in  $p$  rationalizes belief in  $q$ ;

---

<sup>12</sup>Winters [1983], 216.

<sup>13</sup>Wedgwood [2006], 672

<sup>14</sup>*ibid.*, 671

and, perhaps, we could articulate some counterfactual claim about the disposition to the effect that if  $p$  did not rationalize  $q$ , the disposition would not lead from belief in  $p$  to belief in  $q$ . But this would be no different, ultimately, from Winters's account. In both cases, the rationalizing seems to be external to the disposition—a fact about the disposition with no necessary connection to the instantiation of the disposition in the agent. Our reason for calling a manifestation of this particular disposition *inferring* seems to make no contact with the cognitive life of the agent—and so once again it seems accidental that this disposition should play a special cognitive role for her.

The alternative, however, is to understand the disposition as taking two inputs: first, belief that  $p$ , and second, belief that  $p$  rationalizes  $q$ ; and with these two inputs in place, it leads to the belief that  $q$ . Note that this would, in a sense, vindicate the Taking Condition: the taking is here the second belief needed to actualize the disposition. Furthermore, because it relies on a disposition, it seems to circumvent the Carroll-style regress worries we encountered with the doxastic construal of 'taking' considered above. Once again, the *rationalization* does not seem to be doing the right kind of causal work. For rationalization enters in only as another content to which this disposition happens to be sensitive. The causal mechanism is once again no different than that of any disposition of thought—we can no more understand the conclusion as drawn *in virtue of* the rationalization than understand it as drawn *in virtue of* the belief in  $p$ . And so there is no way to understand this disposition to be an *essentially rational disposition*. That a disposition distinguished only by the fact that one of its inputs involves the concept of *rationalization* should be thought to be an essentially rational disposition is just as implausible as the thought that a disposition distinguished by the fact that one of its inputs involves the concept of *a cow* should be thought to be an *essentially bovine disposition*.

More generally, the problem with Dispositionalism about inference is that it cannot distinguish inference, or specially rational dispositions, from non-rational dispositions of thought, except by reference to their accidental properties. Both inferences and, for example, associative tendencies would be causal patterns of thought in precisely the same way. Furthermore, they should, on the Dispositionalist account, be indistinguishable to the thinker. Never does the Dispositionalist rely on the thinker's being aware of the disposition

she is manifesting; she is only aware of the contents of her beliefs. Hence, if she believes  $p$  and subsequently forms a belief in  $q$ , there is no way for her to judge whether the latter belief was formed on the basis of one kind of disposition—inference—or another entirely.

This leaves the status of the belief-formed-by-inferring unclear. For as we remarked at the outset of this investigation, part of what it means to believe a proposition on the basis of inference is to understand the grounds for believing it: to know *why* one believes it, in a sense of that question that asks for a *reason* for belief, not a mere cause. The integrity of the topic of inference relies on the special role inference plays in our conceiving of ourselves as rational; this, in turn, relies on the way that beliefs formed on the basis of inference are, so to speak, self-consciously justified. The Dispositionalist, in solving the causal story of belief-formation at the expense of the Taking Condition, is unable to give an account of inference which lives up to the topic.

So the question remains unsettled. On the one hand, hypostatizing the Taking Condition seems to leave the normativity of the inference inert, alienating it—viciously—from the result of belief formation it ought to have. On the other hand, Dispositionism puts so much conceptual space between the normativity of inference and its causal mechanism that it can no longer account for the exalted status of its object. The goal now, for philosophers writing about inference, is to give an account that can toe the line between these two unsuccessful poles—or that can abandon their framing of the question altogether.<sup>15</sup>

## B. TRANSITION TO THE PROPER TOPIC

It may not be clear what the foregoing has to do with the idea of acting for a reason—my topic here. For one thing, whatever the relation of justification might be in the theoretical case, it does not seem to be shared in the practical case: the fact that Mary is crossing the road does not *justify*—or, more pointedly, *necessitate*—my swerving in anything like the

---

<sup>15</sup>Ulf Hlobil has made some progress in this direction, giving a formal-causal account of the nature of inferring; see [Hlobil \[2016\]](#). Markos Valaris has also provided an account of inference considered synchronically, which I also think holds promise in this regard; see [Valaris \[2017b\]](#). Both of these accounts will be examined in detail the second essay.

same way the fact that  $p$  and the fact that  $p \rightarrow q$  justifies belief that  $q$ .<sup>16</sup> Furthermore, efforts to develop a *practical inference* as an account of acting for a reason have tended to end either in frustration or in constructions that are gerrymandered beyond recognition (or utility).<sup>17</sup> Inference is not the favored apparatus for conceiving of the relation of thought to action, and so, we might think, the problems of inference are simply irrelevant to practical philosophy.

Instead, the topic of acting for a reason centers, naturally enough, around the notion of a *reason*. And indeed, the literature is populated with myriad kinds of reasons: *motivating reasons* and *justifying reasons*; *agents' reasons* and *normative reasons*; *internal reasons* and *external reasons*. Philosophical views of the nature of action and acting for a reason can define themselves by legitimizing one kind of reason while disavowing another; by drawing the line between kinds of reasons in different ways, or in different places; or by giving different accounts of one kind of reason or another. Regardless of the particulars of the theory, the consensus seems to be that understanding reasons is the key to understanding acting.

But I will argue here that, in spite of the shift in terminology, the issues that face reasons theorists on both sides of major debates in metaethics are in fact the same as those which plague the very idea of inference in the theoretical case. The question of acting for a reason is, at bottom, substantially the same as the question of inferring—that is, believing for a reason. Accounts of acting for a reason, as I will demonstrate, face problems exactly analogous to those faced by accounts of theoretical inference. And they err in the same way as the latter erred above—the diagnosis suited to the practical dialectic is, I will argue, the same as that which I gave to the theoretical.

### C. REASONS FOR ACTION AND ACTING FOR A REASON

As I mentioned, there are many kinds of reasons considered by philosophers and many accounts given thereof. In what follows, I will do a case-study, as it were, of two dominant,

---

<sup>16</sup>Though see Dancy [2018] for a thoughtful discussion of this point, and a dissenting view on the matter.

<sup>17</sup>See Dancy [2009] for the former; Broome [2002] for the latter.

and opposing, theories of practical reasons. It is worth noting from the outset that the theories at stake in what follows take themselves to address different topics. The first, Reasons Fundamentalism, gives an account of *normative* reasons—that is, the kind of reason that makes it the case that an action ought to be performed by an agent, regardless of the agent’s comprehension of or attitude toward that reason. The second, the Humean Theory of Reasons, gives an account of *motivating* reasons—the kind of reason that primarily explains why an agent actually performs the action she does.

Because of this way of talking—of one kind of reason as opposed to another—it is easy to suppose that these different kinds of reasons represent two entirely different topics. Treating them, then, under the guise of a single dialectic would be misguided. But as Jonathan Dancy has noted, there are not, really, two sorts of reason; rather, “There are just two questions that we use the single notion of a reason to answer. When I call a reason ‘motivating’, all that I am doing is issuing a reminder that the focus of our attention is on matters of motivation, for the moment. When I call it ‘normative’, again all that I am doing is stressing that we are currently thinking about whether it is a good reason, one that favours acting in the way proposed.”<sup>18</sup> Similarly, as other authors have stressed, the topic of motivating reasons in fact *depends* on the topic of normative reasons to distinguish it from other kinds of “explanatory reasons”—explanations of an action that do not advert to *reasons* in the relevant sense at all. Hence, Sophie might choose to make vegetables for dinner because she thinks they are healthy—her motivating reason involves the healthiness of vegetables and her pursuit of good health. But I might also tell you that she chose vegetables for dinner because I tricked her into thinking vegetables are healthy. That fact does provide an explanation—what some call, somewhat misleadingly in this context, an “explanatory reason”—but it is not her reason for acting, and it does not motivate her—it is not a good reason for her to eat vegetables, nor would she take it to be. The construction of a motivating reason, then, involves some relation to the kind of intelligibility at issue in the concept of a normative reason—motivating reasons involve the ‘speaking-in-favor-of’ that characterizes normative reasons.

The precise relation between motivating and normative reasons is, of course, an open

---

<sup>18</sup>Dancy [2000].

question; different philosophical positions can be staked upon different responses to it. The important thing for our purposes is not to slip too readily into thinking of the topics as separate—or separable. Nonetheless, Dancy is no doubt correct that to speak in terms of these “kinds” of reasons is to address different questions we might ask about the notion of acting for a reason. And it should be stressed, too, that the theories we will presently be examining address themselves to a host of disparate philosophical issues, nearly as many as make contact with their topic.

We will be forgiven, then, for focusing on one question that each theory must address: that is, how one can act for a reason. Insofar as each theory here considered gives an account of reasons-for-action, it is not unreasonable to suppose that sometimes—not all the time, and maybe not for every explicable reason, but *sometimes*—people act for those reasons. Our concern will be how agents relate themselves to reasons, and how action is supposed to be produced as a result.

## 1. Reasons Fundamentalism

We shall begin with the theory recently defended by authors such as Joseph Raz and T. M. Scanlon: Reasons Fundamentalism. Reasons Fundamentalists hold that reasons for action are normative facts, fundamental in the sense that they are not reducible to non-normative facts.<sup>19</sup> A reason is, for Raz, “a fact that actions of a certain kind have properties that can give a point or a purpose to their performance...”<sup>20</sup> Alternatively, for Scanlon, a reason is a kind of normative relation:

“being a reason for” is a four-place relation,  $R(p, x, c, a)$ , holding between a fact  $p$ , an agent  $x$ , a set of conditions  $c$ , and an action or attitude  $a$ . This is the relation that holds just in case  $p$  is a reason for a person  $x$  in situation  $c$  to do or hold  $a$ .<sup>21</sup>

The “reasons” these authors are talking about are, as I mentioned, *normative reasons*. You have no normative reason, in this sense, to drink the petrol even if you believe it is gin—

---

<sup>19</sup>Sometimes “Reasons Fundamentalism” is just taken to name the claim that reasons are the fundamental normative unit, even when an essentially non-normative account of reasons is subsequently given. For the purposes of this chapter, however, I will use “Reasons Fundamentalism” in the more specific sense represented in Scanlon’s and Raz’s views.

<sup>20</sup>Raz [2011], 13.

<sup>21</sup>Scanlon [2014], 31.

even, that is, if we can articulate a motivating reason, or an agent's reason, and explain your acting accordingly thereby. Importantly, these reasons can be correct or incorrect.

The attentive reader will already appreciate the similarity between this conception of a reason and that which was supposed to be the content of the 'taking' invoked by the Taking Condition above. That content made some reference to the relation of support or justification holding between the premises and the conclusion; it articulated, in other words, how the premises were a reason to believe the conclusion. The Reasons Fundamentalists' reason does the same. Saying that a reason-fact states the 'property of an action that gives a point to its performance', or relates that a fact is a reason to do something, is as much to say that it shows the way in which a fact supports, promotes, or justifies the performance of an action.

Reasons, then, are normative facts that hold independently<sup>22</sup> of facts about a particular agent's employing them—just like, we might note, facts about rational support hold regardless of what propositions a thinker happens to believe. But suppose an agent comes to know that some fact is a reason for her to perform an action. How, then, can she act on the basis of such a reason?

**a. Scanlon's account of acting for a reason** For Scanlon, acting for a reason depends on the idea of a rational agent. A rational agent is an agent who can think about and evaluate reasons—these special normative propositions—and, importantly, who is moved by considering reasons. A rational agent, then, if she grasps that  $p$  is a reason to do  $a$ , will do  $a$ . As Scanlon puts it, she will do  $a$  *for this reason*—that is, she will cite  $p$  alone as her reason. Only if "her reference to  $p$  as a reason is challenged"<sup>23</sup> will she cite the more complex fact that  $p$  is a reason to explain herself.

At first blush this seems sensible enough. But on closer inspection it is not obvious what Scanlon means with his terse account of how the rational agent acts for a reason. One possibility is that a rational agent is to be understood as one whose apprehension of certain facts,  $p$ , which are themselves reasons—that is, which stand in the four-place relation

---

<sup>22</sup>Or potentially independently—I address this in detail below.

<sup>23</sup>*ibid.*, 54

detailed above—elicits behavior in accordance with what that fact counts in favor of. The rational agent, then, is one with certain dispositions to action, dispositions actualized by the apprehension of certain facts. But this leaves out of the account any reliance of the rational agent on apprehension of reasons *as such*, that is, the normative relation or relation of support that is captured by Scanlon’s four-place relation. The relation of the agent to the idea of a reason is absent: her disposition is responsive to facts which are, independently and among other things, reasons.

This would be a strange place for a Reasons Fundamentalist to settle: it seems to leave out of the account of acting for a reason the normative relation that is supposed to constitute reasons as such. We should be surprised to find the normativity that distinguishes the Reasons Fundamentalist’s account dropping out of the story when his reasons are taken up by an agent. But we are saved, I think, from so finding, when Scanlon writes that “appeal to the fact that *p* is a reason to *a* may come later, if her reference to *p* as a reason is challenged.”<sup>24</sup> In other words, the more complex fact—the normative relation holding between *p* and the other factors—can be elicited in explanation of the agent’s action. That she can be brought to refer explicitly to the reason-fact shows that some comprehension of this more complex fact is required in order for the rational agent to act for the reason that *p*.

If, then, acting for a reason involves comprehension of this complex proposition, we might begin to worry again about the kind of Sophistication Objection we encountered above. For it seems that we would want to ascribe actions-for-reasons to some agents—children, standardly—who might not yet have the concept of *a reason*. So, for example, it seems perfectly right to describe a child as climbing onto the table because the cookie jar is sitting atop it—indeed, a child could very well explain himself this way. But that does not seem to require that he can talk about that fact *as* his reason—much less that he can articulate anything with the form of a four-place relation. If this is so, then comprehending reasons in Scanlon’s sense is too high a standard for acting for them.

One standard response to this kind of objection is to claim that in acting because the cookie jar is on top of the table, the child reveals an *implicit* grasp of the concept of a

---

<sup>24</sup>*ibid.*, 55.



reason, even if he has not yet been taught the word. But this kind of response undoes the progress we made with this account toward the proper incorporation of the notion of reason. For the child's behavior would be consistent with a disposition that responded to the fact  $p$  alone; and we could choose to ascribe to the child an "implicit" grasp of the concept of a reason if he routinely, on the basis of such a disposition, behaved in accordance with what the fact  $p$  spoke in favor of. In other words, to rely on an implicit grasp of the concept of a reason based on an agent's behavior is to assume that genuine reasons-behavior can occur in the absence of a relation to anything like the irreducible normative fact that the Reasons Fundamentalists was attempting to make central. Acting for reasons, then, if understood this way, is not acting for reasons in the Reasons Fundamentalists's sense at all.

But say we put children once again to the side, and consider agents to whom we can unproblematically ascribe beliefs involving normative concepts. Still it seems that there are multiple ways to understand the account Scanlon means to give of acting for a reason. One possibility is that the rational agent is just one with certain dispositions—not merely to respond to certain facts,  $p$ , alone, but rather to respond to these along with the reasons-facts of which they are constituents. In other words, the Scanlonian rational agent could be understood to have a disposition of the same sort as Wedgwood's "essentially rational disposition" above, which is sensitive to two contents:  $p$  and a content which details  $p$ 's aspect as a reason. But as we saw in that case, so, too, here: the rational agent becomes one distinguished only by the fact that she is responsive to a certain content. When the rational agent forms a belief about a certain relation—the reason-relation—she is moved to act a certain way. The special normative nature of that relation, once again, does no work.

But perhaps Scanlon would not be satisfied with the rational agent understood to be just one among many agents defined more or less arbitrarily by means of contents to which some disposition of theirs is sensitive. In that case, however the rational agent is supposed to be employing her apprehension of the full reason-relation-fact, it could not be so simple as being brutally disposed to respond to it in such-and-such a manner. It seems natural enough, then, to suppose that the process of responding to the reason-relation-fact should involve something akin to reasoning, if it is supposed to be a rational response—putting to work some kind of genuinely rational capacity. So we might suppose that the rational agent must

apprehend  $p$  and the proposition *that  $p$  is a reason for  $x$  to do  $a$  in  $c$* —call this proposition  $R$ —and that she must reason from these two to the performance of the action. But of course in order to reason from the two propositions, she must take them both together to constitute a reason for her to act: she must grasp a reason-relation-fact where both  $p$  and  $R$  together appear as the set of propositions that is the reason—call this proposition  $R_1$ . But of course this  $R_1$  must be grasped in combination with  $R$  and  $p$  as a reason in order for the agent to take herself to have, in fact, a reason to  $a$ —call this further reason  $R_2$ . And so on—we have a familiar regress.

Admittedly, it is not entirely clear what Scanlon’s account of acting for a reason is.<sup>25</sup> But it is clear that working out his account will run him directly into the difficulties that the Taking Condition faced above. On the one hand, there is the threat of a Wedgwood-style dispositionalism that involves recognition of full-fledged reasons-facts, but renders the rationality of the agent and the normativity of the facts which are his ‘reasons’ unrelated—they act in tandem with one another, as it were, but not in concert. But to properly take on the normativity of reasons that is so central to the Reasons Fundamentalist’s doctrine leads to the intellectualist difficulties that the more straight-forward exposition of the Taking Condition introduced.

**b. Raz’s account of acting for a reason** Raz is somewhat more interested in the relation of his reasons to the actions they are so often invoked to explain, in the guise of what he calls the “normative/explanatory *nexus*.” The *nexus* holds that “every normative reason can figure in an explanation of the action for which it is a reason, as a fact that, being recognized for what it is, motivated the agent to perform the action, so that the agent guided its performance in light of that fact.”<sup>26</sup>

But Raz’s account runs into similar issues as we saw with Scanlon’s above, as Doug Lavin has demonstrated. For Raz as well as Scanlon, being motivated and acting for reasons requires the idea of a rational agent exercising her rational powers, so that “the way that the

---

<sup>25</sup>It seems to me that something like the Wedgwood-style dispositionalist account is what is meant by Scanlon; I do not know how he would respond to my objection to it. But since I think that that account is problematic, the principle of charity demanded that I offer an alternative.

<sup>26</sup>Raz [2011], 28

belief [in the reason] has to explain the motivation is by having the content it has, by what it is a belief about.”<sup>27</sup> Part of that content, naturally enough, must be the reason’s “character as a reason.”<sup>28</sup> This would seem to suggest, once again, that acting for a reason involves applying the concept of a reason; once again, then, we find ourselves with a Sophistication Objection. This is what Lavin, in his comments on Raz, calls the first Problem of Intellectualism, namely, that such an account is “psychologically implausible.”<sup>29</sup> Once again, the thought is that there are agents—such as young children—whom we would want to describe as acting for reasons, but to whom it would be implausible to attribute possession of the concept of a reason.

Raz, however, argues that this kind of objection rests on a mistake. The mistake, he claims, lies in the standard for concept possession that is implied by the objection. Raz distinguishes between a high standard and a lower standard for concept possession. The high standard requires being able to explain the concept, “to articulate its implications and inferential connections.”<sup>30</sup> But one can also demonstrate concept possession by merely “follow[ing] the standards for its correct use”—by acting in accordance with the normative requirements a concept places on its bearers. The lower standard allows us to attribute to children concepts like fairness when, for example, they return stolen toys to their rightful owners, or dole out a snack evenly. Similarly, Raz thinks, we can attribute to them possession of the concept of a reason, precisely because we describe them as acting for reasons. In this way, the lower standard for concept possession allows us to attribute the concept of a reason to any agent who might serve as fuel for a sophistication objection.

The problem with this kind of defense is, once again, that it threatens to undo what distinguishes Raz’s theory, as a kind of Reasons Fundamentalism, from a kind of dispositionalism about acting for reasons, or what Lavin calls “automatism.” Such an account attributes acting for a reason to an agent just in case an agent’s behavior in response to a belief manifests a disposition to respond in such a way as has been deemed reasonable. So, observing an agent react to the belief that there is water before her by picking up the

---

<sup>27</sup> *ibid.*, 29

<sup>28</sup> *ibid.*, 29

<sup>29</sup> Lavin [2011], 5

<sup>30</sup> Raz [2011], 32.

glass and drinking, is sufficient for saying that the agent acted for a reason—say, that she was thirsty and drinking water would quench that thirst. Dispositionalism, in other words, will say that an agent acts for a reason just in case her behavioral response instantiates a pattern that is independently—externally, we might say—recognized to be a pattern of reason-reponsiveness. The very thing that goes missing in this account is some sense of—and theoretical reliance on—the fact that the agent takes the reason to be a reason to act. Raz’s loose conception of concept-possession threatens to turn his account into just another form of Dispositionalism. But Reasons Fundamentalism, of the sort Raz is aiming to defend, is meant to put recognition of reasons *as reasons* at the fore in explaining action.

So the Sophistication Objection afflicts Raz, too. But let us once again set aside cases where an extended sense of concept-possession would need to be invoked. Still, Raz’s account runs into the further problems we saw with Scanlon’s. As we have seen, the Reasons Fundamentalist requires that an agent recognize the fact that  $p$  that is her reason to *be a reason* in order to act for the reason that  $p$ . This means, consequently, that belief in  $p$  alone is not sufficient to amount to comprehension of a reason, and acting because  $p$  alone not sufficient to amount to acting for a reason. But neither, of course, is comprehension of the fact *that  $p$  is a reason* sufficient to rationalize an action—the agent must know that  $p$  is the case. So it seems the agent must have two beliefs: the belief that  $p$ , and the belief that  $p$  is a reason to  $A$ ; and, since knowing either of those propositions alone did not suffice to be a reason for her to act, then they must be combined in order to be a reason. But in order for the Reasons Fundamentalist to recognize the agent as acting for a reason, the agent must comprehend her reason—now her complete reason, the combination—to be a reason. Hence the agent must also have a belief that the conjunction of  $p$  and the fact that  $p$  is a reason to  $A$  is a reason for her to  $A$ . And this last belief, of course, must be combined with her belief that  $p$  and her belief that  $p$  is a reason. Once again, we have a regress.<sup>31</sup> Hence, Raz’s account, too, falls prey to the objections that plagued the Taking Condition above.

Reasons Fundamentalism attempts to make normativity an entirely external matter, embedded in certain special facts that we call reasons. But if the normative force of reasons

---

<sup>31</sup>Lavin calls this the problem of ‘Contemplativism’ (Lavin [2011], 9), the second horn of the dilemma he presents for Raz’s account; it is ‘contemplative’ because it presents the agent as locked, so to speak, in the theoretical reasoning behind acting, never able to get to doing anything.

is entirely external to the agent, then that agent has little recourse to relate herself to it. On the one hand, she may be related cognitively to the normativity of her reason. But just as an apprehension of a taking-fact proves to be inert in accounts of theoretical inference, so, too, apprehension of a reasons-fact appears to be inert in the case of acting. On the other hand, if she does not make cognitive contact with the reason, then the normative force simply remains external to the agent; she may be tuned to react to the fact that is a reason in a certain way, but the normative force of that reason appears not to play a role in determining her behavior.

If making normative force external to the agent seems hopeless, then rendering it in some sense internal would seem to be a plausible alternative. So it is to that alternative that we now turn.

## 2. The Humean Theory of Reasons

On the opposite side of the debate about reasons we find a kind of Humeanism about reasons. The Humean Theory of Reasons, as we shall call it,<sup>32</sup> claims that having a reason to do A requires having a desire that would be promoted by doing A. The argument for the Humean Theory of Reasons has two main components. First, one must accept the thought that a reason that is incapable of motivating an agent is no reason at all—only something that could motivate an agent can be called a reason. This is the claim of Reasons Internalism.<sup>33</sup> Second, one must accept that agents are only ever motivated to act because of the presence of some desire. This is what is claimed by the Humean Theory of Motivation. If one is only able to be motivated because of the presence of some desire; and if having a reason requires being able to be motivated; it follows that having a reason requires the presence of some desire.<sup>34</sup> The Humean Theory of Reasons is, then, a theory of motivating reasons—but it implies that motivating reasons are the only kind of reasons there are.

To be more precise, let me give Michael Smith's definition of a motivating reason:

R at t constitutes a motivating reason of agent A to  $\phi$  iff there is some  $\psi$  such that R at t consists of a desire of A to  $\psi$  and a belief that were he to  $\phi$  he would  $\psi$ . (36)

---

<sup>32</sup>Following Mark Schroeder.

<sup>33</sup>Schroeder [2007], 200. The *locus classicus* of course is Williams [1979].

<sup>34</sup>See again Schroeder [2007].

It is clear enough what it is to have the kind of means-end belief involved in having R; what it means to desire to  $\psi$  is less clear. But Smith helpfully provides an account. Desires, Smith claims, are states with a certain functional role—namely, grounding dispositions to act: “like the disposition to  $\phi$  in conditions C...”<sup>35</sup> Hence the desire in R should be understood to be grounding A’s disposition to  $\psi$  in some conditions.

Now presumably by adverting to this particular motivating reason we would be aiming to understand A’s act of  $\phi$ ing. And it seems that for the purpose of this example we should understand the desire to  $\psi$ —evidently a complex action relying on  $\phi$  as a means—to ground the disposition to take the means to  $\psi$ . Now, the disposition to take the means to  $\psi$  seems to be actualized by the belief that  $\phi$  is the means to  $\psi$  with the result that the agent  $\phi$ s. Hence we have a picture like the following: what it means to have the desire to  $\psi$  is to have a disposition to  $x$  in conditions that include recognizing  $x$  to be a means to  $\psi$ . The agent acquires the belief that  $\phi$ -ing is a means to  $\psi$ -ing; her disposition is actualized, and she  $\phi$ s. Since the reason is, for the Humean, the combination of the disposition to take the means to  $\psi$  and the means-end belief that actualizes that disposition, the agent acts for—and because of—her reason.

It should be obvious how the Humean Theory is analogous to Dispositionalism about inference—Smith’s account of desire renders the Humean account itself a form of Dispositionalism. The Humean identifies the ‘reason’ as the whole mechanism, in a sense: both the disposition to respond (that is the desire) and the belief to which the disposition is responsive. Because they understand having a reason ultimately in terms of being in a causally efficacious state—that is, having certain dispositions to act—Humeans seem to have no problem connecting their reasons to the actions they are meant to produce. The problem is that it is unclear whether the Humean can give any grounding to the idea that what is produced is a genuine *action* at all.

For consider: having a disposition does not require knowing that one has it, much less that one has some conception of its content. Hence being disposed to take the means to  $\psi$  does not require that the agent know he is so disposed, nor that he recognize himself to have any disposition towards  $\psi$  at all. Indeed, Smith takes this to be a virtue of his account—

---

<sup>35</sup>Smith [1987], 52.

understanding desires to be dispositions allows us to be fallible with respect to our desires, as common sense would deem us sometimes to be, because dispositions do not come along with any necessary apprehension of them.<sup>36</sup> So, not knowing that I am in any way inclined toward  $\psi$ , I acquire the belief that  $\phi$  is a means to  $\psi$ ; since I in fact have the desire to  $\psi$ , straightaway my disposition to take the means to  $\psi$  is activated and I begin to  $\phi$ .

If a passerby were to ask me why I am  $\phi$ -ing, it seems that I would be hard-pressed to answer. As far as I can tell, I formed an errant belief about some means to some end, and then suddenly found myself propelled into motion. But if my answer to the question ‘why are you  $\phi$ -ing’ is either ‘am I?’ or ‘my, I don’t know!’, then it is hard to see how I could be described as acting for a reason. What’s worse, if we follow Anscombe in thinking that intentional actions are those to which a certain sense of the question ‘Why?’ is applicable, then we cannot say that I am  $\phi$ -ing intentionally at all. And clearly, my relation to the act of  $\psi$ -ing is even more attenuated.

The Humean account of acting for a reason, then, is no account of *acting* at all. The dispositional analysis the Humean gives of desire as the motivating force behind action seems no different in kind from any other disposition to movement a human being might have, which will include compulsions whose effects we would by no means wish to categorize as intentional actions, or as performed for a reason. The sense in which the Humean agent acts for a *reason*, in other words, has gone missing; no rationality, no thought is called for. The Humean might object that the actualizing belief that  $\phi$  is a means to  $\psi$  is being given too short-shrift in this understanding; but I think the burden of proof would be on the Humean to show why this content should do special work. Since the idea of a *means* is showing up merely as a content of a belief, it seems to be on a par with any other content; and surely dispositions to behavior could be actualized by other contents without our wishing to attribute to the agent anything like an intention—as when, for example, I begin salivating at the mention of a very special breakfast pastry.

Just as Dispostionalism about inference failed, in the final analysis, to distinguish between rational and non-rational dispositions of thought, so too Humeanism about reasons seems unable to distinguish intentional actions from mere behavior or, so to speak, animal

---

<sup>36</sup> *ibid.*, 53.

movement. Dispositionalism about inference left the conclusion belief of unclear status—in need, as Boghossian pointed out, of rational assessment. Humeanism about reasons leaves the status of the action *as* action unclear—the account it provides does not suffice to ensure that an action was performed for a reason in any recognizable sense, and so it fails to ensure that the agent truly acts. Just as Dispositionalism ultimately seemed to jettison the topic of inference, it becomes impossible to understand Humeanism to be a theory of reasons at all.

#### D. ON THE VALUE OF THE ANALOGY

I have tried to develop an analogy between a dialectic in the literature on theoretical inference and two poles in the literature on acting for reasons. The latter is afflicted in the same manner as the former by the tension between rationality and causation. On one side we find an account that over-intellectualizes what is required by apprehension of a reason and thereby alienates that reason from the action it is supposed to elicit; on the other we find an account that is too brutally causal to accommodate the special rational status of its effects.

The analogy, I think, suggests a way to understand the accounts given of acting for a reason as, more fundamentally, competing conceptions of the nature of a rational process *as such*. In the theoretical case, of course, the process results in belief; in the practical case, action. If we follow Elizabeth Anscombe’s method of identifying the topic of a philosophical account of acting, then we would have independent reason to conceive of the topic as that of the nature of a rational process. Anscombe, as I have mentioned, identifies intentional actions as those “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.”<sup>37</sup> But this is as much to say that actions can be picked out purely by their being a certain kind of rational product; we can rely on their relation to rationality to define the realm of discourse about them.

The reasons-for-action literature does not seem to be concerned either with so abstract

---

<sup>37</sup>Anscombe [1957], 9.



or with so formal a question as I have settled on here. Indeed, it can come to seem that two substantive conceptions of the human agent are at stake—one driven by reason, the other slave to the passions. Nevertheless there is reason to think that the question of a rational process as such is primary, even for these theorists. For it is not as straightforward as it might seem to distinguish Reasons Fundamentalism and the Humean Theory of Reasons on the lines of the classical debate concerning the ascendancy of reason over desire or of desire over reason. Reasons Fundamentalism, for its part, does not rule out a desire-theory of reasons, where an agent has a reason to do A just in case doing A will satisfy one of her desires. Desire could find its way into one or another of the variables in Scanlon’s reason-relation: either in the articulation of the circumstances  $c$ , or in the specification of the nature of the agent  $s$ . Given Scanlon’s conception of a reason, it could turn out that reasons-facts hold only when the agent has a certain desire. Indeed, Scanlon himself recognizes a certain kind of desire-theory as compatible with his Reasons Fundamentalism.<sup>38</sup> It is similarly open to Raz to claim that reasons-facts are always related to desire-facts. The Reasons Fundamentalist’s machinery could always rely on desires, then, in the content of reasons.

Similarly, the account Humean theories give of ‘desire’ seems abstract enough to accommodate whatever the Reasons Fundamentalist would want to rely on in explaining acting for a reason. Take, for example, Davidson’s description of the pro-attitude that serves as the desire in his belief-desire theory: among the things to be included under that heading are “desires, wantings, urgings, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values...”<sup>39</sup> Or take Bernard Williams’s description of the ‘subjective motivational set’ whose involvement in the constitution of reasons marks internalism: “ $S$  can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may abstractly be called, embodying commitments of the agent.”<sup>40</sup> These conceptions of what can play the desire-role in an account of action do not limit the nature of the motivation-mechanism very much. Nor should they—the more restrictively ‘desire’

---

<sup>38</sup>Scanlon [2014], 5.

<sup>39</sup>Davidson [1980], 4.

<sup>40</sup>Quoted in *Internal Reasons: Contemporary Readings*, ed. Kieran Setiya and Hille Paakunainen (2011), 41.

is characterized, the less plausible the account becomes. But note that the more abstractly ‘desire’ is conceived, the less the account is distinguished from its supposed rival. Smith’s Humean understands desires as dispositions to act in certain circumstances, just so sparsely conceived. But Scanlon’s account of the rational agent could, as I argued, conform to just such a dispositional analysis. So Scanlon’s rational agent could be one defined, in the terms favored by the Humean, by the presence of a certain desire.

This is not to deny that there are substantive debates to be had between the Reasons Fundamentalist and the Humean about reasons. But it is, I hope, suggestive that those debates are better not put in terms of substantive claims about psychology. As I have tried to show, they are better understood as competing accounts of what has been revealed as the practical analogue of inferring: namely, acting.

This last aspect of the analogy we should pause to reflect on. From the perspective of this paper, *acting*—and now I shall assume that all acting, in the sense in which we are interested, is acting for a reason<sup>41</sup>—is analogous to *believing on the basis of inference*, or *inferring*. This suggests that acting could come to be regarded as on equal footing with inference so far as the life of the rational animal is concerned. We have, that is, no reason as yet to assume that the problem of theoretical inferring is prior to that of acting; and indeed, insofar as both inferring and acting appear to be in need of the same kind of account, we have reason to see them as coeval.

Furthermore, our analogy affords us a new purchase on the Aristotelian Thesis—the thesis that the conclusion of the practical inference is an action. I mentioned in passing earlier the disfavor into which the idea of the practical inference has fallen. With no account of the practical inference yet given—and no consensus on its nature forthcoming in the literature—we might wonder why we should even introduce the topic. But it is worth noting here that objections to the idea of the practical inference generally proceed under the assumption that the theoretical inference is unproblematic.<sup>42</sup> My survey of the literature on the theoretical

---

<sup>41</sup>Dancy usefully distinguishes ‘acting for a reason’ and ‘acting in light of reasoning’ (Dancy [2018]); I am assuming that the action that the philosopher is interested in is acting for a reason, but not necessarily acting in light of reasoning. This is as much to say that the philosopher is primarily interested in intentional action—perhaps as the perfection of the kind (see Anton Ford’s essay in Ford et al. [2011]). This assumption furthermore rules out mere behavior, reflexes, and other things that might be thought to be voluntary, but are not, again, full-blooded actions in the philosopher’s sense.

<sup>42</sup>See, for example, Broome [2002] and Paul [2013].

inference has shown that this is not so. And I have tried to bring out the ways in which the nature of inferring is in need of explanation. Acting, it turns out, is in need of precisely the same explanation—accounts of it are subject to the same pitfalls as accounts of inferring. So whatever we might mean by a practical inference (naturally, an account must be given, and I have done no work toward that here), it seems to me that what is meant by introducing the idea is that acting—just like believing—is the product of a rational process. So I take the Aristotelian Thesis to be vindicated to this extent: the value of the practical inference is in directing our mind toward the nature of action as the product of a specially rational process, and so there is no question of a practical inference apart from conceiving of it as concluding in an action.<sup>43</sup>

## E. CONCLUSION

I have argued that the nature of acting for a reason can be understood as the nature of a rational process as such, in the same way that theoretical inference can be understood as a paradigmatic rational act. I have argued for this claim by showing what theoretical inference and acting for reasons have in common—or rather, by showing that both are afflicted by the same serious problems. In the next chapter, my aim is more positive: I turn to solutions to the problem of the theoretical inference, and a corresponding theory of action which, I argue, is poised to solve the problem of acting for reasons.

---

<sup>43</sup>Here, as everywhere, we must be careful to set our sights on the properly delimited topic. Obviously we can reason about practical things, have inferences to other attitudes (like intention), go through hypothetical practical inferences which do not conclude in action. But the central topic, I am claiming, is the rational production of action.

### III. A NON-STANDARD APPROACH TO INFERENCE AND ACTING FOR REASONS

#### A. INTRODUCTION

This essay concerns the question: “how do we act for reasons?” Put this way, however, no answers even begin to suggest themselves: the question itself mystifies. We might be tempted to precisify by noting that there are reasons—the fact that I promised, say—and there are actions—mowing the neighbor’s lawn; the question would then seem to be, “how do we get from one to the other—from reason to action?” But that sort of question sets visions of Rube Goldberg machines dancing in our heads. As will become clear, I think that approaching the question in this way inevitably leads to error. Instead, I should like to address the question, “What do we mean when we say that we act for reasons?” or perhaps, “We say that we act for reasons—what is the *for* relation to which we refer?”

But I will be addressing this question by a roundabout path. I want, first, to examine the topic of the theoretical inference—that is, one answer to the question, “how do we believe for reasons?” Theoretical inference is, in other words, supposed to be one way a thinker might form some belief on the rational or logical basis of other beliefs. But authors investigating the nature of inference face a dilemma. Intuitively, in order for the thinker rationally to form a belief that Q on the basis of P, she must in some way take it that P supports Q and form the belief that Q because of this fact. But this intuitive condition—which, following Paul Boghossian, we will call the Taking Condition<sup>1</sup>—lands us on one horn of the dilemma. The most natural way to satisfy the condition is to posit a state—a belief, for example—the

---

<sup>1</sup>See [Boghossian \[2014\]](#).

content of which details P’s rationally supporting Q, as a necessary part of an inference.<sup>2</sup> But, as theorists have noted, a state with this content—in particular, involving the concept of rational support—appears to be too high a cognitive load for some thinkers whom we take, nonetheless, to be inferers. Furthermore, even if we grant that all inferers can grasp such contents, it is not at all clear what role that state is supposed to play in the drawing of the inference—and vicious regresses and vicious circularities await the obvious answers.<sup>3</sup>

To avoid impalement, some philosophers have rejected the intuitive Taking Condition, and offered in its stead various dispositional accounts of inference. According to Dispositionism, roughly, A infers Q from other beliefs P only if P and Q instantiate an inference form and A’s transition from P to Q is the result of a general disposition to make transitions that exhibit that form. Whereas accounts formulated in accordance with the Taking Condition appeared to provide no way for the transition from premises to conclusion to occur, Dispositionism puts the transition front and center, relying on what is assumed to be a familiar notion of causation to account for how it is effected. But Dispositionism faces its own difficulties. A dispositional account leaves open what the precise causal mechanism is that underlies the disposition in question. This means that any number of causal mechanisms might underly a disposition exhibiting the requisite form. But it seems that at least some of these will seem to us to be deviant: though the mechanism leads from the right cause to the right effect, it nonetheless proceeds along a path that would undermine the integrity of the inference. (Consider, for example, dispositions that rely on the intrusions of a mad scientist—and the like.) But worse, by making the transition between premises and conclusion causally determined, Dispositionism seems to leave out the idea of understanding the transition as itself genuinely rational. Inferences, according to this objection, appear to be on a par with all other mere dispositions of thought—at any rate, so far as the inferer herself is concerned. But then we might well wonder whether Dispositionism is an account of *inference* at all.

Hence, the dilemma: on the one hand, accounts of inference aimed at satisfying the

---

<sup>2</sup>There may, of course, be other ways to satisfy the Taking Condition—as we shall see. Boghossian himself favors a rule-following view. But in general, objections to the Taking Condition involve conceiving of its satisfaction in roughly this way.

<sup>3</sup>See, for the former, Carroll [1895]; for the latter, Boghossian’s arguments against the “background condition” approach in Boghossian [2014].

Taking Condition appear to over-intellectualize the act of inferring, and seem to leave the fact that, in inferring, a belief is formed unexplained; on the other hand, accounts that jettison the Taking Condition in favor of a causalist program do not seem intellectual enough: belief-formation seems to be only accidentally related to the rationalization-relation at the inference's core.

Recently, however, there has been a spate of solutions proposed to this dilemma: accounts of inference that aim to vindicate the Taking Condition while giving a satisfactory account of how inference is to be effected. These accounts often depart radically from the traditional framework from which theorists have approached the question of inference—and from within which the dilemma has arisen. And though they are all distinct from one another, they are the more striking, I think, for what they share in common.

I will begin this essay by presenting two of these new accounts of inference. They are, of course, worth examining in their own right, and I will treat them accordingly. But I do not intend to come to any conclusions about the nature of inference or the state of the debate on the theoretical inference. I want to argue, instead, for a certain conception of intentional action—the “non-standard story” of action developed by theorists impressed by Elizabeth Anscombe’s work on intentional action—by means of this dialectic in the literature on the theoretical inference. In particular, I will argue that the new, radical conceptions of inference being proposed to solve the above dilemma bear are analogous to a particular conception of the nature of intentional action. Furthermore, as it has been argued elsewhere,<sup>4</sup> the problems afflicting our understanding of the nature of the theoretical inference arise as well for our understanding of how we can act for reasons. If, then, the non-standard, Anscombean conception of action stands to the “standard story” of action—the causal theory of action—as the newly proposed solutions to the problem of inference stand to the traditional conception of inference, then that should, I claim, count in favor of the Anscombean conception of action.

---

<sup>4</sup>See [Lavin \[2011\]](#), as well as Chapter 1 of this dissertation.

## B. THE NON-STANDARD ACCOUNTS OF INFERRING

### 1. Hlobil’s Force Account

Ulf Hlobil’s is one of the voices raised in support of Boghossian’s Taking Condition. The Taking Condition, recall, states: inferring necessarily involves the thinker’s taking her premises to support her conclusion and drawing her conclusion because of that fact. Hlobil prefers to focus his discussion of inference on what he calls “Inferential Moorean Absurdity”—the absurdity that arises from the utterance of sentences like,

(IMA)  $P$ ; therefore,  $Q$ . But the inference from  $P$  to  $Q$  is not a good inference.<sup>5</sup>

Understanding utterances like (IMA) as absurd appears to imply something like the Taking Condition: if adding the belief that your inference is no good to your stock of beliefs clashes (in some way) with the rest of your mental life, then you must (in some sense) take your inference to be good. Thus, if we think that sentences like (IMA) are absurd, then we must affirm the Taking Condition.

Yet, as Hlobil recognizes, working out what the *taking* in the Condition is supposed to be is notoriously difficult. Hlobil himself discusses and rejects both beliefs and intuitions as candidates for takings, on both familiar and unfamiliar grounds.<sup>6</sup> In response to these problems, Hlobil proposes a solution that eschews the need to analyze takings as one or another familiar kind of state—intuitions, beliefs, seemings, etc.—entirely. Instead, Hlobil proposes that we think of the taking involved in the Taking Condition as analogous to the attaching of doxastic force by which Frege understands the act of judgment.

For Frege, to judge is to “acknowledg[e] a thought as true”<sup>7</sup>—or in other words, to take some thought to be true. But that taking-to-be-true is not an additional content added to or combined with the content of the thought; it is, rather, the subject’s attitude toward the content of the thought. And it is precisely in bearing this attitude—this relation—to the thought that the subject becomes rationally responsible for the truth of the thought. So, in

---

<sup>5</sup>See [Hlobil, 2014, 421]. I have reversed the letters here in order to cohere with my own favored usage.

<sup>6</sup>His “unfamiliar” argument concerns takings as beliefs, and depends on the idea that takings cannot be transmitted by testimony. See Hlobil [forthcoming].

<sup>7</sup>Quoted in [Hlobil, forthcoming, 17].

bearing this relation, the subject cannot also come to believe that the thought is not true, without being, by her own lights, irrational.

Taking-to-be-true thus characterizes the *force* of the judgment that P: when I judge that P, I attach this particular kind of force—“doxastic force”—to the thought that P. Hlobil proposes that we understand inferring analogously: to infer Q from P, I engage in the act of attaching *inferential force* to some thoughts. The thoughts that make up the content of an inference—that is, the thoughts to which I attach inferential force—are just the premises and conclusion, in the form of a structured set, called an “argument.”

But what is inferential force? Just as doxastic force is a relation one bears to a proposition that can be described as ‘taking-P-to-be-true,’ inferential force is the relation one bears to an argument that can be described as “taking-P-to-support-Q.” Once again, this ‘taking-P-to-support-Q’ is not the content of a judgment; it is the relation the thinker bears to the contents P and Q when she infers Q from P.

Furthermore, Hlobil claims, attaching inferential force to an argument has consequences for the thinker’s attaching of doxastic force to the relevant propositions. As Hlobil puts it, “it is metaphysically impossible to attach doxastic force to the premises of an argument to which one attaches inferential force and also fail to attach doxastic force to the conclusion of the argument.”<sup>8</sup> Which is to say, if a thinker attaches inferential force to the argument P, therefore Q; and she attaches doxastic force to P, then, as a matter of metaphysical necessity, she must attach doxastic force to Q (if she does not give up attaching doxastic force to P, that is).

So, in sum, inferring is the act of attaching inferential force to an argument. Notice, first, how Hlobil’s account vindicates the Taking Condition—in particular, through the idea of Inferential Moorean Absurdity. To reiterate, in the case of judgment, judging that P is in rational tension with denying that P is true, even without the fact that P is true appearing in the content of the judgment that P, because taking-P-to-be-true describes the relation of the thinker to the thought. So, in the case of inferring, inferring that Q from P is in rational tension with believing that this is a bad inference—believing, for example, that P does not support Q—because ‘taking-P-to-support-Q’ describes the relation of the thinker to

---

<sup>8</sup>[Hlobil, forthcoming, 20].



the argument. Thus, sentences like (IMA) are absurd because of the nature of the attitude involved in inferring Q from P; and thus, the Taking Condition is vindicated.

But the Condition is vindicated without having to posit a separate propositional attitude with a content that concerns rational support. The very idea of support characterizes the thinker's *attitude*, not the content of her thought. Thus, sophistication objections, as well as Carroll-style regresses, cannot get a foothold in this view. Thus, Hlobil maintains the Taking Condition—and indeed, its intuitive support, in the shape of Inferential Moorean Absurdity—without over-intellectualizing his conception of inference.

## 2. Valaris's Semantic Account

Markos Valaris, meanwhile, takes a different tack. Valaris, like Hlobil, is impressed by the Taking Condition, which he thinks captures the sense in which reasoning is something “*we do*.”<sup>9</sup> Since reasoning is, as Valaris emphasizes, an active, personal-level process—the process of, as he puts it, “*making up our own minds*”<sup>10</sup>—it is something for which we bear responsibility. And because we bear responsibility for our reasoning, we are legitimately criticizable for it. This is, he thinks, captured by the Taking Condition: the Condition ensures that reasoning reflects “*the subject's take* on what her evidence requires.”<sup>11</sup>

But Valaris is, of course, sensitive to the problems that arise for the Taking Condition. His solution is to reject the presupposition that inferring must be understood as a causal process—that is, the process of some beliefs causing the drawing of a conclusion. According to Valaris, it is this conception, and not the idea that the thinker must take her premises to support her conclusion, that has caused problems for the Taking Condition.<sup>12</sup>

To understand Valaris's account, we begin with a simple thought: the premises and conclusion of an inference have meaning, which must be understood by the thinker who is reasoning with them. What is it to understand a proposition? According to Valaris, “[i]ntuitively, understanding a statement involves *knowing how it represents things as being*,

---

<sup>9</sup>[Valaris, 2017a, 4].

<sup>10</sup>*ibid.*, 4.

<sup>11</sup>*ibid.*, 5.

<sup>12</sup>See Valaris [2014] and Valaris [2017b].

or *what things have to be like for it to be true.*"<sup>13</sup> Valaris suggests we analyze the content of this knowledge in terms of "possibilities, or ways for things to be."<sup>14</sup> Thus, understanding the sentence, "this dog is a border collie" involves understanding which possibilities are ruled out by that sentence's being true: namely, all possibilities in which this dog is some other breed; all possibilities in which the dog's parents were not border collies; etc. We should note, however, that understanding the sentence "this dog is a border collie" need not involve the capacity to describe all the possibilities relevant to the truth of the sentence—things involving, say, the genetic make-up of this dog and his predecessors. Instead, as Valaris puts it, it involves the capacity to "*pick out* the relevant possibilities upon considering them."<sup>15</sup>

Coming to believe a statement, then, involves ruling out the possibilities that are incompatible with the statement's being true. If I come to believe that this dog is a border collie, I thereby rule out any possibilities in which this dog is a chihuahua or a spaniel; I also thereby rule out any possibilities in which this dog's parents were chihuahuas or spaniels; and so on.

What has this to do with reasoning? From the perspective of this possibilities-analysis of the meaning of statements, we can say that Q follows from P just in case there are no real possibilities in which P is true and Q is not true.<sup>16</sup> Thus believing that Q follows from P involves recognizing that you have already ruled out all possibilities in which Q is not true, simply by believing that P. But notice: if you reflect on your belief that P and realize that all possibilities in which Q is not true are also ruled out, you *thereby* believe that Q. In other words: recognizing that the possibilities available to you to believe all involve the truth of Q *is* just believing that Q.

This, then, is Valaris's account of reasoning: recognizing that Q follows from P, when P is something to which you are already committed, *just is* believing that Q. Or in other

---

<sup>13</sup>[Valaris, 2017a, 14].

<sup>14</sup>*ibid.*, 14. Note here that of course there are other ways of understanding understanding or belief, but I will not consider them here.

<sup>15</sup>*ibid.*, 15. According to Valaris, this is a cognitive skill, but not one that is constituted by further reasoning—an important note given his criticism of rule-following accounts as taking for granted reasoning in the application of the rules of reason, and thus as circular. We might worry that this, too, asks too much of a thinker in understanding a sentence: after all, if presented with various dogs' genomes, I could not pick out the border collie's. But I think Valaris's conception of belief could be satisfied if not all relevant possibilities, but some sufficient number of those that are, as it were, epistemically relevant—suitably related to the rest of one's beliefs—could be picked out by the believer.

<sup>16</sup>*ibid.*, 16.

words: the *taking* involved in the Taking Condition *just is* the drawing of the conclusion of an inference. Valaris’s account thus abandons the idea that taking one’s premises to support one’s conclusion somehow *causes* acceptance of the conclusion—instead, taking one’s premises to support one’s conclusion simply *constitutes* acceptance of that conclusion.<sup>17</sup>

### 3. Lessons Learnt

Hlobil’s and Valaris’s accounts are by no means the same—indeed, the two authors object to one another’s theories in writing.<sup>18</sup> But I think that what is similar between the two accounts is more striking than what is distinct. In this section, I want to draw our attention to the features that these—and other<sup>19</sup>—novel accounts of inference share. Even if an individual account proves untenable, I think it should impress us that multiple philosophers are looking in the same direction for an answer to what has seemed an intractable problem. Indeed, I think it is suggestive that that direction holds some real promise.

**a. From causal process to act** The first thing that we should note about Hlobil’s and Valaris’s accounts of inference is that they both move away from a picture of inferring according to which it is a *causal process*, and toward a conception according to which it is a special kind of *act*. The traditional conception of inference assumed that inferring was a causal process that brought a thinker from premise-beliefs to a concluding belief: the goal was to describe the premise-beliefs as causing “in the right way” the drawing of the conclusion. But both Hlobil and Valaris aim to move us away from such a model.

For Hlobil, this involves understanding inference as the act of attaching a certain kind of force to an argument—that act, then, bearing metaphysical implications for the thinker’s doxastic state. In this case, the entire argument is unified in the single act of attaching inferential force, rather than parts of the argument playing different causal roles in the production of another, separate, conclusion.

---

<sup>17</sup>[Valaris, 2014, 13].

<sup>18</sup>Hlobil argues against Valaris’s belief-based account in Hlobil [forthcoming]; Valaris argues against some of Hlobil’s arguments for diachronic norms of rationality in Valaris [2017b].

<sup>19</sup>See, for example, Rödl [2013] and Kimhi [2018]. Both of these authors provide accounts of reason or reasoning that share the important aspects of promising accounts of inference that I outline below.

Valaris, meanwhile, is quite explicit about rejecting the causal-process model of inference: “reasoning...is not to be *identified* with a causal process.”<sup>20</sup> According to Valaris, inferring is instead to be understood as the act of recognizing that some proposition follows from what you believe—which is one way of coming to believe that proposition. Whereas the causal conception would place the premises, and the recognition of their entailment relations, at a distance from the conclusion, and then connect them by means of the disputed causal relation, Valaris instead leaves no room between the elements of the inference, uniting them in the act of taking the premises to support the conclusion.

Moving away from understanding inference as a kind of causal process allows us to disentangle questions of justification from questions of causation—the entanglement of which was one plausible culprit in generating the problem of inference in the first place. According to the standard framework, the problem of inference is precisely the problem of identifying a relation between premises and conclusion which is both justificatory and, *in virtue of being justificatory*, also causal. It is both the description of a goal, and the description of why that goal must be frustrated: justification and the kind of causation that haunts these accounts appear to be by their nature irreconcilable. Hlobil and Valaris, by characterizing inference not as a causal process, but instead as an act, allow us to set aside this whole, ill-fated problematic.

But surely, one might object, acts also occur in the causal order—why shouldn’t the problem of causation arise as well for accounts that understand inference as a special kind of act? In other words, if the causal etiology of the concluding belief proved problematic, why shouldn’t the same hold for the causal etiology of the act of inferring itself?

We can admit that there might well be questions about the causal etiology of the act of inferring. But on reflection I think that it is clear that these questions will not generate skepticism about the rationality of the inference in the way that understanding inference itself as a causal process did. Asking after the causal etiology of the act of inferring would be like asking “what made you realize that you wanted to be an astronaut?” The answer might be something like, “I saw a shooting star, and then I realized”—yet we feel no pressure to think that the star is itself the real *reason* for the decision, nor do we think that a story

---

<sup>20</sup>[Valaris, 2014, 29].

like this impugns the rationality of making such a decision. So, too, what precedes the inferring in the causal order does not seem to threaten the inference’s rational status—whereas understanding the inference itself as causally determined does.

But recall the Taking Condition: inferring necessarily involves the thinker’s taking her premises to support her conclusion and drawing her conclusion because of that fact. The Taking Condition appears to require that we understand the relation between the taking—recognizing that your premises support your conclusion—and the concluding precisely as causal. Don’t Hlobil and Valaris thus fail to satisfy the Taking Condition, since they abandon conceiving of inferring itself in causal terms? I do not think so—instead, these authors force us to reimagine the nature of the causal relation that holds. Hlobil, for his part, is explicit in saying so: “the ‘because’ means that the taking is the formal cause of the inference; it is not among its efficient causes.”<sup>21</sup> Formal causation, of course, describes the “form” or the “account of the what-it-is-to-be” of a thing—to say that X is the formal cause of Y is to say that Y is what it is because of X; X makes Y what it is. For Hlobil, taking-P-to-support-Q, as we have seen, characterizes the nature of the act that the thinker brings to bear on the argument; that act constitutes the argument as an inference for the thinker. Thus, the taking—the special force characterized as the taking—makes the inference what it is: it is its formal cause.

Valaris does not use the language of formal causation, but he might as well have: he writes repeatedly that the taking—that is, the belief about support—plays a “*constitutive*, rather than *causal*, role in reasoning.”<sup>22</sup> As we have seen, the act of coming to recognize that some of your beliefs imply another belief *just is*, for Valaris, coming to have that belief; in other words, recognizing the entailment-relation amounts to forming the concluding belief—and indeed, makes that believing the drawing of an inference. Thus, for Valaris, too, the taking is the formal cause of the inference.<sup>23</sup>

In this way, even though these authors eschew the idea that inference must be understood

---

<sup>21</sup>[Hlobil, 2016, 2].

<sup>22</sup>[Valaris, 2014, 4].

<sup>23</sup>I should note that something’s being a formal cause does not preclude it from at the same time being an efficient cause. In this case, however, I mean to vindicate the idea that we can speak of a “because” precisely without committing ourselves to the existence of some efficient-causal chain. Whether some such chain exists, I leave an open question; my argument is, rather, that an analysis of inference in terms of such a chain is hopeless.

in efficient-causal terms, they vindicate a formal-causal framework in which to understand the role that the taking plays in inferring. Formal causation, meanwhile, appears not to generate any of the difficulties that the efficient-causal model of the standard story did.

**b. From atomism to complexity** Lewis Carroll's Tortoise challenges Achilles to force him logically to accept the conclusion of a certain inference. As Achilles proceeds, he asks the Tortoise to accept, one by one, additional premises in what now threatens to be an infinite inference. Each time the Tortoise accepts a premise, he has Achilles write it down in his notebook—as a sign of the Tortoise's acceptance.

As Valaris argues, the traditional conception of inference appears to take this picture—that acceptance of a premise is akin to writing it down in a notebook—quite seriously.<sup>24</sup> Acceptance is all one kind of act, the same no matter what its object: belief. And inferring is simply the moving from some beliefs to some other belief. Indeed, all of our cognitive capacities (at any rate, in this context) seem to be constructed out of the belief-relation, that special attitude we bear toward certain propositions.

Inference, then, has precious little material with which to be understood: there is belief, invariant across beliefs, and then there are its contents. As theorists have noted, the contents bear all kinds of interesting logical relations to each other; but that does not necessarily transfer to the beliefs. And the beliefs, meanwhile, have no internal signs to distinguish them from one another—they have different contents, but otherwise they are identical: same notebook, same pen.

Both Hlobil and Valaris jettison this model of inference, in different ways—they both aim to introduce acts with a more complex internal structure into the picture to give an account of inference. For Hlobil, this act is precisely the act of inferring: the attaching of taking-P-to-support-Q force. This force not only takes as its object a structured set of propositions, rather than just a single proposition; but it also is, by its own nature, a more complex kind of act. For Hlobil, the characterization of inferential force itself involves taking one thing to support another; even as an act, then, it has a kind of internal structure to it. Furthermore, for Hlobil, the act of attaching inferential force has metaphysical implications

---

<sup>24</sup>See Valaris [2017b].

concerning other acts: inferential force therefore also reaches outside of itself, connecting it to the larger web of the thinker's mental life. This, too, can be understood as a kind of complexity of the act itself: it is, by its nature, bound up with a set of other acts. Neither of these things can be said for the traditional Achilles's-notebook-conception of belief: it is atomic and, as such, appeared to be isolable.

Valaris, for his part, calls our attention to the materials made available simply by a richer conception of belief or acceptance. Belief, Valaris reminds us, is not as simple as writing a proposition down in a mental notebook. Rather, to believe a proposition involves ruling out (epistemically) certain possibilities for oneself, and ruling in other possibilities. Thus, believing—as the act of managing the possibilities available to one—is itself an internally complex set of acts. Furthermore, given that believing P and believing Q may involve ruling out quite different sets of possibilities, and given that believing is just the act of ruling out these possibilities, it seems as though all beliefs are distinct, and are rendered distinct by what their objects are. Every belief, in other words, involves a distinct act of partitioning, or distinct set of acts of ruling out and ruling in possibilities. And we can note, as well, that the object of belief—as involving a possibility-space and the partitioning of that space—is itself internally complex: the content of a belief is not at all like a sentence written in a notebook.

The move away from an atomic conception of belief—both as regards the act itself, and as regards its content—to a more complex conception both of the acts involved in inferring and of the contents of those acts, provides both Hlobil and Valaris with a much richer set of materials with which to give an account of inference. If, as it seems, the problem of inference is rooted in—or at any rate, related to—the limited resources called upon in the traditional account (Achilles's notebook and pen), then introducing some kind of complexity into the account of acts and contents seems essential to solving the problem.

**c. An act that contains its own grounds** If elucidating inference in terms of complex, non-atomic attitudes is essential to showing how we can infer, it also makes room for the idea that, in drawing the conclusion of an inference, the thinker's act contains its own grounds.

On the traditional conception of inference, what it means to infer is to move *from* some

propositions and *to* another proposition. This can, indeed, seem essential to the question of inference: in examining inference, it seems as though we are examining a special kind of belief-formation process; thus, the result of this process must be, simply, a belief. And, as we have seen, according to the traditional conception, beliefs are invariant in their nature—no matter how they are formed.

But once we conceive of the relation between the premises and the conclusion in this way, the familiar philosophical dialectic comes into view. For, in doing so, we place the conclusion at a distance from the premises; thus we must find some other element to connect them. The Taking Condition aims to connect premises and conclusion by means of a claim about the relation of support between them; but, given the traditional conception of the materials available for the account, this claim can only play the role of another content, and, as such, can no more connect premises and conclusion than the premises themselves could. Dispositionalism connects premises and conclusions by means of that trustiest of adhesives: a *causal connection*. But causal connections—holding, as they do, between things like billiard balls and the like—cannot do justice to the *rational* nature of the connection between premises and conclusion. And so, once sundered, premises and conclusion are lost to one another.<sup>25</sup>

Both Hlobil and Valaris aim to bring premises and conclusion together again by rendering them inseparable in the very conception of inference. For both authors, inference is an act that by its nature involves both premises and conclusion; it is not a process that leads from one to the other. According to Hlobil, inferring itself is the attaching of the special inferential force to *both* premises and conclusion—*unifying* them all under the guise of a single act. Inferring in this way has implications for the simpler act of judgment, but it is not itself those implications—it is not the passing on to that judgment. Rather, inferring is precisely the act in which the premises and conclusion are brought together. Thus, in inferring, the thinker does not pass along to something separable from its grounds; by inferring the thinker

---

<sup>25</sup>Again, my invective against causation, here, should be understood as focused on the thought that causation can provide an analysis or account of inferring. I recognize that proponents of the causal theory will simply insist that the kind of causal chain that constitutes an inference is *the right kind* of causal chain, allowing us to set aside deviant causal chain worries and the rest. But, like Davidson, I despair of specifying what it means for the premises to cause the conclusion *in the right way* without making reference to concepts like justification and the like, which suggests to me, anyway, that *causation* is no longer playing a real role in the account.



does not leave her premises behind.

Similarly, for Valaris, the act of concluding the inference just is the act of recognizing one's grounds as grounds. In other words, when I believe some proposition, I rule out certain epistemic possibilities; when I realize that all the possibilities still open to me involve some other state of affairs, I thereby—in that very act of recognizing—come to believe that proposition. Or, put another way, in recognizing that I have grounds for some belief, I recognize that I have that belief—that is, I believe it. Thus for Valaris, too, there is no distance between the premises of an inference, and its conclusion; accepting the premises as premises *is* accepting the conclusion. Here, too, the state that results from an inference is inseparable from the inferring itself: both are simply a special cognizance of some aspect of my epistemic possibility-space. And so, again, in concluding an inference, I never leave my premises behind, or move on to a separable belief-state; rather, my acceptance of my conclusion contains within it my acceptance of the grounds.

This is, I think, the most important aspect of these new accounts of inference—and we can see how the other elements I have emphasized are essential to understanding this point. Inference is not a process—it does not lead to a separable conclusion. Instead, it is a single, complex kind of act, one which precisely unifies premises and conclusion. Separating premises from conclusion renders the special status of inference—the special status of this kind of belief-formation process—unintelligible. However we are to understand inferring, its product—its result, or endpoint—cannot be thought to leave behind its origin. The conclusion of an inference, however we are to understand it, must in some way contain its own grounds.

### C. A NON-STANDARD STORY OF ACTION

Thus far, I have presented two new accounts of the nature of theoretical inference, and I have attempted to draw from them general lessons about the conceptions that must be left behind and the materials that must be put in place in order to solve the problem of inference. In this section, I aim to show that a similar movement is possible—and, indeed, necessary—when

thinking about action.

As has been argued elsewhere, the idea of acting for a reason is afflicted with a dilemma analogous to that facing theorists of inference. If, then, solutions to the problem of inference have a certain shape, it is not unreasonable to expect solutions to the problem of acting for a reason to have a similar shape. What is striking, I think, is that a certain conception of the nature of action—one not motivated by this particular problem—already shares the important features of solutions to the problem of inference that I identified above. Thus, if we think that the problem of acting for a reason is a serious one—as I do—and we can show that one conception of the nature of intentional action is particularly suited to solve it—as I intend to do—then we will, by these means, have an argument in favor of that particular conception of the nature of intentional action. Providing that argument is, then, the over-arching aim of this essay.

To begin, I will briefly present the “standard story” of action, which will play a role analogous to that played by the traditional conception of inference above. Then I will present the alternative account, emphasizing the features it shares in common with the novel accounts of inference.

## 1. The Standard Story of Action

The so-called “standard story of action”—that is, the one that appears to enjoy dominance among philosophers of action—is some version of the Causal Theory of Action, which we can formulate as follows:

Any behavioral event A of an agent S is an action if and only if S’s A-ing is caused in the right way and causally explained by some appropriate nonactional mental item(s) that mediate or constitute S’s reasons for A-ing.<sup>26</sup>

Strictly speaking, the Causal Theory of Action is not one single theory, but a family of theories, individuated by distinct specifications of either what amounts to an A-ing being caused *in the right way*, or what particular nonactional mental items do that causing. For our purposes here, however, a general sense of the Causal Theory will suffice; I will discuss,

---

<sup>26</sup>[Aguilar and Buckareff, 2010, 1].

for simplicity's sake, a toy version of Donald Davidson's causal theory.<sup>27</sup>

Here is the standard story of how an action is born, according to the standard story: An agent finds herself with a desire—to turn on the light, say—and a belief that by flipping this switch, she will achieve illumination. Together, this desire and this belief amount to a reason for the agent to flip the switch. This reason, then—the belief-desire pair—causes the agent to move her finger thus-and-so against the switch. That finger-movement can also be described as a turning on of a light; the latter is a description under which the action is intentional because the agent had reason to do it and, once again, the event that constitutes it was caused in the right way by that reason. (I will, with Davidson, refrain from giving an account of in what this *right way* consists.)

That should be enough about the standard story of action for our purposes: an action is an event—a bodily movement—that is caused in the right way by certain special mental antecedents, which also bear a justificatory relation to the descriptions under which we take the event to be an intentional action. It is constituted as an action in virtue of being so caused.

## 2. The Non-Standard Story in Three Movements

The alternative to the standard story is, like the standard story, really a family of theories, all influenced by the philosophy of G. E. M. Anscombe. In what follows, I will be developing this story using Anscombe's own writing and what it seems to me (and others) follows from what Anscombe has herself said on the subject. In doing so, I hope to isolate features of the Anscombean conception of action that all or most theorists who take themselves to be offering a non-standard story of action would accept.

**a. First movement: rejection of a necessary causal relation between reason and action** Our non-standard story begins with a rejection of the claim that an account of action can be given by identifying the special causal etiology of certain events. The idea

---

<sup>27</sup>Davidson is supposed to be the contemporary father of the Causal Theory, but if that is so, then he is a filicidal one: he himself provides the 'wayward causal chains' objection to the Causal Theory, and opines that the problem is "insurmountable." See [Davidson, 1980, 79]. More on this below.

that an account of action must make reference to the fact that reasons are also *causes* of actions appears to be irresistible to proponents of the standard story because of what is known as “Davidson’s Challenge.”<sup>28</sup> As Davidson puts it, if we do not claim that reasons are causes, then “something essential has certainly been left out, for a person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it.”<sup>29</sup> So, for example, suppose that I have two reasons to turn on the light: first, the fact that I want to read a book, and second, the fact that there is a prowler outside (and I’ve spotted him) and I want to scare him off. As it happens, I turn on the light only for one of these reasons—the latter, more pressing reason, say. In virtue of what is it true that I turned on the light in order to scare off the prowler and not in order to read a book? According to the causal theorists, the answer must be: in virtue of the fact that *that* reason, and not the other, caused my action.

But it seems to me that the Challenge does not, in fact, force us to speak of causes. For it seems to me that it is possible to say that the referent of “the action” changes throughout Davidson’s terse argument: “a person can have a reason for *an action*, and perform *the action*, and yet this reason not be the reason why he did *it*.” As Davidson, of course, recognizes, talk of reasons for actions introduces a “quasi-intensional” context, in which it matters tremendously what description of the action is given.<sup>30</sup> So, the fact that I want to read a book may well be a reason to turn on the light—but, if turning on the light in fact startled my cat, it is nonetheless wrong to say that the fact that I wanted to read a book was a reason to startle my cat. On the other hand, multiple descriptions of an action might well be felicitously related to a given reason: the fact that I want to read a book is a reason to turn on the light; in this case, turning on the light involves flipping the switch, so the fact that I want to read is a reason to flip the switch; and so on.

Now take the case at hand: the fact that I want to read is a reason to flip the switch, turn on the light, and thereby, say, make reading light available. The fact that I want to scare off the prowler is a reason to flip the switch, turn on the light, and thereby alert the prowler to the fact that I am home. Davidson’s challenge requires us to say: I have a reason—the

---

<sup>28</sup>[Aguilar and Buckareff, 2010, 7].

<sup>29</sup>*ibid.*, 9.

<sup>30</sup>[Davidson, 1980, 5].

fact that I want to read—to turn on the light, since to do so would be making reading light available; and I turn on the light—and thereby, we must maintain, make reading light available; but I do not turn on the light for that reason. But as should be clear, if I do not turn on the light because I want to read, then, in turning on the light, I am not *making reading light available* (intentionally). And this is because *making reading light available* is not rationalized by my desire to scare off the prowler; indeed, a *different* action is: namely, alerting the prowler to the fact that I am home.

In other words, Davidson’s Challenge depends on taking two actions that share a description (or two!) in common to be the same, when they are not. An action performed for one reason will be, in virtue of that fact, different from an action performed for another reason.<sup>31</sup> And so, even though we might advert to causes to explain the difference between turning on the light for one reason and turning on the light for another, we do not *have* to: we can instead simply examine more closely the action performed.

Davidson himself, it seems to me, plants the seed of the thought that an account of action need not make reference to the cause of action. In “Actions, Reasons, and Causes” he writes that an action will be caused *in the right way* by its reason, and seems to think that this ‘right way’ could be given a non-circular account; but later, in “Freedom to Act,” he begins to “despair of spelling out...the way in which attitudes must cause actions if they are to rationalize the action.”<sup>32</sup> And there, and later still in “Psychology as Philosophy,” he conjectures that specifying ‘the right way’ for the attitudes to cause the action would involve in some way “a chain or process of reasoning that meets standards of rationality.”<sup>33</sup> This is, of course, just a suggestion of his, and not a worked-out account. But if something like this is right—that specifying the right kind of cause will have to make reference to practical reasoning—then it seems to me that the concept of a *cause* is left doing no work in the account. That is to say, if the idea of a causal link between a reason and an action does not get us to action on the basis of that reason, but rather must be further specified by reference to practical reasoning, then it seems as though the causal link is *not* what is ‘left out’ in Davidson’s Challenge and is, therefore, not to be given pride of place in our account.

---

<sup>31</sup>I will return to this point in great detail in the third essay.

<sup>32</sup>[Davidson, 1980, 79].

<sup>33</sup>[Davidson, 1980, 232].

Rejecting the claim that actions *must* be explained by reasons as causes allows Anscombeans to appreciate the wide variety of ways that actions are, in fact, explained. Some authors have focused particularly on teleological explanation as the primary explanation of action: the agent did A *in order to* do B.<sup>34</sup> Other authors have focused on explanation of action by action, or explanation of action by re-description of the action. As Anscombe points out, a perfectly suitable response to the question “Why are you A-ing?” is, “I am B-ing.” Michael Thompson even claims that this form of action-explanation—what he calls, “naive action-explanation,” explanation of an action by the “progress of the deed itself”<sup>35</sup>—is the primary form, and that “sophisticated” forms of action-explanation—explanations adverting to precisely the aspects the Causal Theory takes to be essential—are in some sense derivative.<sup>36</sup>

For our purposes here, what is striking is the tendency of authors, once freed from the grips of the standard story, to turn to the action itself in some form or another for explanation. Moving away from a theory that places causation by independent states at its center, authors focus instead on the action itself, and seek explanation in the elements that are, as it were, internal to the thing needing to be explained. This allows authors both to set aside the causalist traps that threaten any “process-based” account, but it also forces them to scrutinize more seriously the nature of *action*. In this way, the Anscombean theory resembles the new accounts of inferring: they too move away from the idea of inference as a causal process; they embrace, instead, a conception on which a certain special act of mind is at issue. And in doing so, they, too, examine more carefully the materials provided by that special act itself.

**b. Second movement: focus on the internal complexity of actions** The non-standard story, in turning toward the actions themselves, begins with a recognition of the internal complexity of a typical intentional action. To start, most actions are temporally extended—potentially very widely: writing a paper might take a month; orchestrating a

---

<sup>34</sup>See, for example, [Sehon \[1997\]](#).

<sup>35</sup>[\[Thompson, 2008, 90\]](#).

<sup>36</sup>See [Thompson \[2008\]](#). There are still other forms of explanation of action that neither mention the agent’s mental states nor the action itself. So, for example, there are the anti-psychologistic arguments made by [Dancy \[2000\]](#), and, indeed, represented in the theory of reasons presented in [Scanlon \[2014\]](#).

coup, a year. And such actions have internal parts, which are themselves actions. First I sow dissatisfaction among the ranks; then I take over the airwaves (this too will take some time, and has parts); and so on. Taking over the airwaves is a means to orchestrating a successful coup; but it also may be the way that I am orchestrating a coup at some given moment. That is to say, *in* taking over the airwaves, I *am* orchestrating the coup; either description is a correct answer to the question, “What are you doing?”

Furthermore, the parts of an action are not indifferently related, but have an important structure to them. Suppose, to take a more straightforward example (for there are many ways to pull off a coup, I think), I am making an omelet. The parts of making an omelet include cracking eggs, whipping them up, pouring them into a pan, and so on. But it matters in what order the parts of the action are performed: first you must crack the eggs—indeed, this can be understood as a means to whipping them, when whipping them is a means to making an omelet; if you whip them before cracking them, then not only are you not making an omelet (at any rate, not successfully), but it seems that your whipping is something else entirely from the whipping involved in my omelet-making. So the parts of an action are organized by the idea of that action’s going well.

Most of our actions have parts in this way. But are there not also point-like actions, which are, therefore, not internally complex? Take, for example, turning on the light; or shooting someone. Both are nearly instantaneous; what are their parts? Of course, in truth, such actions are not instantaneous, but are temporally extended like the rest of our actions. And if something goes wrong—the bulb is dead, or the trigger catches—then it will become clear that these actions, too, have parts. But even in the happy case, a kind of complexity is exhibited. For my act of turning on the light can be given multiple other descriptions: flicking a light-switch; preparing to read a book. All of these will be accurate descriptions of the same act: that little, almost instantaneous finger-movement. And these descriptions, too, are not indifferently related, but have the same means-ends structure built into them: I am preparing to read a book by turning on the light; I am turning on the light by flicking the switch; but I am not flicking the switch by preparing to read a book. Thus even an action that seems point-like can be seen to have a complex internal structure.

In this way, Anscombeans bring to our attention—remind us of, really—the fact that

actions themselves are not homogenous or atomic, but are instead internally variegated and complex. And, just as theorists of the theoretical inference, rejecting the idea that inference must comprise atomic beliefs toward atomic contents, are able to enrich the resources for their accounts; so, too, Anscombeans seem poised to be able to give a richer account of the relation of reasons to actions.

**c. Third movement: actions contain their own grounds** Implicit in all this discussion is the most important aspect of the Anscombean conception of action: actions contain their own grounds; in conceiving of actions, we never leave reasons behind.

To begin with, for Anscombe, calling something an intentional action at all “has reference to a *form* of description of events.” She goes on, “What is essential to this form is displayed by the results of our enquiries into the question ‘Why?’”<sup>37</sup> In other words, for Anscombe, a description of an action is *formally* the description of an intentional action, which is to say, something that is subject to the question ‘Why?’ in a sense that asks after a reason for acting. So, something’s being described as an action at all makes reference to the fact that it is performed for reasons; being an action is determined essentially by being reasoned in this way.

But the point goes beyond the idea that actions *in general* are necessarily performed for reasons *in general*. In fact, a particular intentional action is defined—it is determined, or made the action that it is—by the particular reasons for which it is performed. And therefore in that way, too, an action is inseparable from its reasons.

To bring this out, let us suppose that I am making that omelet. I crack the eggs; my reason for cracking the eggs is that I am making an omelet. But ‘making an omelet’ is also a description of what I am doing *in* cracking the eggs: at this moment, what my omelet-making consists in is my cracking the eggs. My cracking the eggs *is* my making the omelet here and now, or is a *stage* of my omelet-making. And my egg-cracking is a part of omelet-making because of the reason for which I am cracking the eggs: namely, the fact that I am making an omelet. Hence, the reason for performing the action—cracking eggs—makes that action what it is: a part of omelet-making, or an egg-cracking that is an omelet-making.

---

<sup>37</sup>[Anscombe, 1957, §47, p. 84].



According to Anscombeans, these kinds of sentences—that ‘cracking these eggs is, here and now, my making an omelet’—are not mere eccentricities of language, but instead express something important about the individuation of actions. An egg-cracking that is an omelet-making is a different action than an egg-cracking that is a cake-making, and, still more obviously, both of these are different actions than an egg-cracking that is, say, a bomb-making (if there is such a recipe). What an action is a part of determines what that action *is*; often, what the action is a part of will be, at the same time, the reason for which the action is performed; hence the reason for which an action is performed determines what that action is. An action, in the very way in which it is individuated, does not leave its reasons behind: those reasons constitute the action as the action that it is.

The reason-structure that relates ends to means as reasons also unifies the means, or parts, of an action into a whole. That is, it is in virtue of the structure given to omelet-making by practical reasoning that the parts of omelet-making—cracking eggs, whipping them up, etc.—are held together in a whole: the act of making an omelet. It is the fact that I have set out to make an omelet that determines that I should crack the eggs; it therefore determines that this egg-cracking is a part of omelet-making; and so, when the egg-cracking is done, and the egg-whipping, and so on—each of these parts, in turn, being defined by the fact that they are a part of making an omelet—the whole amounts to my action of making an omelet. Reasons unite a series of events into a whole action: it is on account of the structure that reasons provide that we are able to have single, unified, temporally extended actions, as opposed to mere happy series of events.

The relation of the parts of an action to the whole—that is, the way in which the parts are individuated by means of that relation, constituted as an omelet-making in progress, for example—this relation is what is uncovered by the question ‘Why?’ As Anscombe shows, repeated application of this question can uncover the whole internal structure of an action—here, most clearly perhaps, in the case where we feel inclined to speak of many descriptions of the same, as when I turn on the light. This structure we can call, following Anscombe, the ‘A-D order’ internal to the action: I am doing A in order to do B; I am doing B in order to do C; I am doing C in order to do D. My answers to the question ‘Why?’—that is, my reasons for acting—will also provide a structured set of descriptions of what I am doing: I

am doing D by doing C; I am doing C by doing B; I am doing B by doing A. The A-D order, then, precisely captures the identity of my reasons with my action.

The relation of the whole to its parts—the way in which the parts are unified into a whole that can be thought of as a single action—this relation, meanwhile, is what is exhibited in practical reasoning, prospectively, from the first-person perspective. In practical reasoning, we have an aim—for example, making an omelet—and we set about answering the question ‘How?’ Here, too, we will discover in our thought the A-D order, but now in reverse: I begin by asking how I might D, and conclude that I must C; in order to C, it transpires, I must B; until finally I arrive at the decision to A. In answer to the question ‘How?’—that is, as the content of our practical reasoning—we find the D-A order. Thus, again, our action is revealed to be constituted by our reasoning about it.

Actions, then, never leave their reasons behind. There is not a reason, on the one hand, and an action, on the other, that might or might not issue from this reason. The reasons for a particular action determine what that action is, how it unfolds, what its parts are, and how it is to be understood as a completed whole. Reasons, then, are internal to understanding an action as the action that it is; you cannot lay hold of an action without at the same time laying hold of its reasons.

This argument is not yet fully general. There are two kinds of cases that seem to elude it—that is, actions about which I have not yet established that they are characterized by their reasons. To see the first, take some action—the pumping of water into a house—about which we can elicit the string of descriptions A-D: the agent As in order to B; Bs in order to C; and Cs in order to D. One possibility, which Anscombe brings to our attention, is that not all of the action-descriptions in such a chain are actually being effected: D, that is, might be so lofty a goal that it seems incorrect to say that here and now, by means of the act of pumping, the agent is D-ing. In this case, though B-ing is constituted by its reason—that the agent is C-ing—it seems to be false, strictly, to say that C-ing is constituted by its reasons, namely some act of D-ing—because Ding is not, in fact occurring. Sometimes, to use Anscombe’s phrase, the reason for which I am performing the action will be beyond “the break” ([Anscombe, 1957, §23]). But in these kinds of cases—when the action-description which serves as the reason is not in fact being effected—I think my point can be made with

the small expedient of the notion of *trying*. After all, though it may not be true that the agent is D-ing, it will, in the case I have described, be true that she is trying to D. And something's being an attempt is often as central to its identity as its being a successful action. And so, I claim, C-ing may still be understood and constituted by its reason: namely, that the agent is trying to D.

The second kind of case is D itself—whether or not it is being effected. Suppose, for ease of exposition, that it is. And suppose that, if the agent is asked “and why are you D-ing?” she provides, not another, wider action-description, but a different sort of answer—for example, what Anscombe calls a “desirability characterization.” So, for example, she says that she is D-ing “because it is right” or “because it is my duty.” These kinds of answers are also, as it were, ‘beyond the break’: they abandon the explanation of action by action and turn to another form. But these kinds of reasons, too, it can be seen, effect the identity of the action, and so can be said to constitute the action in the way I rely on here. Take two seemingly identical acts of exercise: two agents run five miles. One does it ‘for her health;’ the other ‘because her life-coach told her to.’ It seems to me that these acts—though identical in all “internal” respects (they both turn to the right in order to leave the driveway; they both leave the driveway in order to get on the five-mile-loop, etc.), can nonetheless be understood to be different actions: one, for example, is an expression of the virtue of health; the other is a performance of duty. And it should be emphasized that this difference is not merely in intention or other (supposedly) purely mental aspects, detachable from the act itself, but rather is a difference in the action. It is neither metaphorical nor, I think, elliptical to say that some actions are themselves expression of virtue, and that others are performances of duty. But these descriptions of the action are precisely given by reasons that come after the break. So here, too, the action is constituted by its reasons; here, too, the action does not leave its reasons behind.

## D. PROSPECTS OF THE NON-STANDARD ACCOUNTS

So far, I have presented new, non-standard accounts of the nature of inference, followed by the Anscombean non-standard story of action, and I have tried to show the ways in which all of these non-standard accounts are similar. All of them reject the idea that their object is analyzable as a certain sort of causal process by which reasons (or premises) cause a thereby-justified result; instead, they emphasize the fact that the so-called ‘result’ is not even so much as characterizable unless we conceive of its reasons as, in some sense, internal to it. Drawing a conclusion, in the theoretical case, is not the same as any given mental-state formation; the act itself is characterized by having the rational-support structure built into it. Similarly, acting is not like any other kind of movement, but is, rather, essentially so as to be for reasons.

But are these accounts promising—or do they represent philosophical wrong-turns to be avoided? I have argued that they do a good job of addressing the problems introduced by the traditional accounts; but the possibility remains that they introduce intractable problems themselves.

Let us consider the theoretical case first. It seems to me that the obvious worry about the non-standard accounts of inference will have to do with the individuation of the concluding-beliefs. Hlobil’s account, by positing the inferring act as *distinct from* but *metaphysically necessitating* the formation of certain belief states may avoid this problem: given the metaphysical connection, Hlobil’s account will result in a separable belief-state that can be given whatever characterization philosophers choose to give it. Meanwhile, there would seem to be little objection to the inferring act itself being individuated by means of its premises and conclusion.

One might object to Hlobil’s account that it is not particularly parsimonious: it introduces both a new primitive act of mind—attaching inferential force—and basic metaphysical connections between acts. Furthermore, both of these posits appear, from a certain perspective, *ad hoc*: they can almost seem to be acts introduced by giving names to the open questions about inference. But I myself am never too sure how much stock to put in objections like these—I’ve never had the taste for desert landscapes myself—and so I will simply

set them aside.

Valaris's account, on the other hand, does seem susceptible to the problem of belief-individuation. Recall that according to Valaris's account, if I have some belief, P, and then I realize that Q follows from P, I thereby believe that Q: recognizing that all the epistemic possibilities open to me from the perspective of my belief in P leaves only possibilities in which Q is true simply amounts to believing Q. My beliefs, in other words, are internally related to one another: my belief that Q, since I arrived at it by means of an inference from P, is importantly determined by that belief.

But now suppose some other thinker believes R; it turns out that in all the epistemic possibilities open to this thinker, Q is also true; realizing this, the thinker thereby believes that Q. What we want to be able to say, of course, is that this thinker and myself have both formed the same belief: namely, Q. But it seems that on Valaris's account, we cannot say this. My belief in Q is controlled by a belief in P; it involves only possibilities in which P is true as well; but the other thinker's belief is controlled by R; and it seems perfectly possible that the possibilities in which R and Q are true are a different set than those in which P and Q are true. So it seems as though the contents of our two beliefs in Q are different: different epistemic possibilities comprise each of the beliefs. And therefore it seems as though we do not *share* any belief at all.

The objections, in other words, to Valaris's account would be the traditional objections to holism about belief. And I am not in a position to render judgment about that debate here. But I will make a few observations. First, it seems to me that there is a lot going for this kind of holism about belief. Imagine a thinker who says that she believes that whales are mammals. But suppose this thinker will not assert or assent to any related claims, such as that whales are not fish; whales give birth to live young; and so on. At some point, I think, we will no longer feel able to attribute to this thinker the belief that whales are mammals; in the face of this, we might say, our belief-ascription "falls to the ground." And so it is clear that, to some extent at least, knowledge of the inferential relations a proposition bears do in fact constitute belief in that proposition. That is, they determine what we *can say* about the beliefs of a given thinker.

I think, furthermore, that there are a number of bullets that we easily could—and, indeed,

perhaps should—bite in order to maintain holism about belief. Is it really so bad that we cannot say that both Cervantes and I believed the same thing in believing that grass is green? If we take a really distant thinker, then this thought seems downright attractive: I think that we *should* believe that my belief that grass is green and a primitive Mesopotamian's belief that grass is green are simply distinct. Meanwhile, in cases where the thinker seems to be more of a peer—Newton, say—why not simply admit that our beliefs are, indeed, distinct, but only in very minor ways that we can ignore in ordinary contexts? After all, I take it that what is irksome about the idea that Newton's and my beliefs that grass is green are not the same is that, in ordinary contexts, it is irresistible to say that they are. And yet, it does not seem to me as though the concept of *belief* appears very often in ordinary contexts: in natural language, it appears to mark out either a case in which someone takes something to be true that is known to be false (belief in Santa Clause), or when someone believes something without sufficient, or at any rate scientifically standard, evidence (belief in God). Neither of these uses characterizes the kind of belief that philosophers are interested in (at least in this context); which suggests that the latter kind of belief is technical and philosophical from the start. In which case it seems not at all threatening that we shouldn't be able to ascribe to me and Newton the same belief: it seems perspicuous, and, indeed, philosophically illuminating.

Again, I do not intend my discussion of holism about belief to be either comprehensive or conclusive. More important for our purposes here is the question: does a similar problem arise in the practical case? For just as in the theoretical case we end up with highly specifically individuated beliefs—beliefs that bear the stamp of their grounds on their contents—so, too, in the practical case do we end up with actions that are highly specifically individuated: actions which are given their identity by all of the various reasons for which they are performed.

Thus if two agents attend the union meeting for different reasons, it is false to say that they did *the same thing* for different reasons; in fact, they did two different things: one showed support for the union's cause, and the other put into motion a plot of sabotage. Or again, two agents mow their neighbors' lawns: one does so because he promised; the other does so to get his nagging wife off his back. Here, too, we cannot say that they did the same

thing: one kept a promise, and the other capitulated to his wife. Here, too, there are two different acts corresponding to the agents' different reasons.

But surely this is correct: surely we do want to distinguish the acts of these agents. In other words, right off the bat it seems to be none but a good thing to individuate actions this finely: it seems to give us a clearer picture of what, in fact, is going on. Thus whatever we find intuitively irksome in the case of finely-individuated beliefs does not seem to bother us in the practical case—we find ourselves instead simply with accurate descriptions.

If it is important to us to say that, nonetheless, our lawn-mowers did *something* in common—namely, mow the lawn—I think the Anscombean account is perfectly capable of doing so. According to the Anscombean, these two actions do share *a description* in common. And this is, of course, important, since the descriptions under which an action is intentional are those that we care about in talking about action. But as I have tried to bring out, we can be more perspicuous about an action by considering the set of descriptions under which it is intentional, and indeed, by considering the structure that holds together this set. Thus, whereas our two actions do, indeed, share a description in common, the full story reveals that the actions themselves are not, after all, the same.

But this conception of actions as finely-individuated is not only amenable to our ordinary way of thinking—it is, in fact, philosophically useful. The Anscombean story of action is in a position to solve a variety of problems in ethics that arise as long as the idea of action is left unquestioned. Indeed, it seems to me that certain problems in ethics arise precisely because the causal theory of action leaves the internal nature of actions themselves without content. Only given the Anscombean theory according to which actions contain and are individuated by their grounds can we solve such problems.

Objections to the Doctrine of Double Effect provide one clear example. Stephen Makin has convincingly argued that such objections rest on a failure adequately to recognize the fact that the descriptions under which an action is intentional—and so, permissible or impermissible—depend on the agent's intentions and thus her reasons. The Doctrine of Double Effect claims, roughly, that it is sometimes permissible to bring about as the consequence of one's action what it would be impermissible to intend directly. The Doctrine is supposed to be compelling as an explanation of cases like the terror bomber/strategic

bomber pair: whereas it seems to be morally wrong to bomb and kill 100 civilians in order to terrorize the population of a hostile government into submission, it seems to be permissible to bomb a munitions factory in order to weaken the opposing military, even if it is known beforehand that 100 civilians are employed in the factory and will, during the bombing, be killed. The moral difference between the two bombers, according to the Doctrine of Double Effect, is produced by their reasons for acting.

Authors like T. M. Scanlon and Judith Jarvis Thompson claim to find it mysterious that the reasons for which an agent acts could make any moral difference. Whether a given action is permissible, they claim, should only have to do with the action—and its consequences; the agent’s reasons might have something to do with the agent’s goodness or badness, but surely nothing to do with the permissibility of the action itself.

As Makin points out, however, if we recognize that “actions are prohibited as falling under certain descriptions” and that “the intention of the agent determines the nature of the action—that is: what the correct description of the action *is*”<sup>38</sup>—or as we might put it, the agent’s reasons determine the nature of the action—then it follows straightforwardly that the agent’s reasons are relevant to the permissibility of the action. In other words, recognizing that there is no thought of action at all apart from thought about reasons—as the Anscombean account brings out—shows us that any moral question about actions (with permissibility or prohibition as paradigm such questions) is at the same time a question about reasons. Thus philosophical mystification in the face of what appears to be an intuitive ethical doctrine can be dispelled by embracing the Anscombean conception of action and the principle of action-individuation that it implies.

A similar point can be made about the topic of moral worth. Suppose two agents donate equal sums to charity—one does so for the sake of the children, the other does so in order to improve her (bad—and accurate) reputation. According to some philosophers, these actions are equally morally good, but they are different in terms of their moral worth. And so the question is: how can two performances of the same action have different moral qualities? Traditionally, an action’s moral worth has been thought to depend on the reasons for which the action is performed. But neither the quality nor the strength of reasons appears to track

---

<sup>38</sup>[Makin, 2017, 8-9].



the moral worth of an action in just the right way. So, it seems, the puzzle of this moral difference remains.

If, however, we locate the reasons within the action itself—if we come to see the action as a more complex particular bearing the mark of all its reasons—then we can come to see moral worth as a measure of all the different aspects of the given action itself. The question, then, is not how the same action (a donation to charity) might have different moral qualities; it is rather a question, once again, about two *different* actions.<sup>39</sup> Here, too, embracing the Anscombean conception of action can turn what appears to be a philosophical puzzle—the puzzle of moral worth—into an interesting ethical implication of our action-theory.

I suspect that more puzzles in ethics may be resolved in just the same fashion: that by appreciating the inner complexity of actions themselves—rather than taking ‘action’ for granted and focusing instead on reasons, intentions, or other mental states—we can elucidate the issues that arise when actions appear in ethical theory. And this is, I think, a noteworthy result. It is especially so when we consider the fact that fine-grained belief-individuation is, if not problematic, at least unintuitive—in light of that fact, it is striking that in the practical case, fine-grained individuation of action appears only to count in favor of the Anscombean theory.

If the Anscombean theory provides a plausible and philosophically productive account of action, then the fact that it is suited to avoid the dilemma that faces the idea of acting for a reason provides, I think, very good reason to accept it over the standard story.

## E. MORALS OF THE (NON-STANDARD) STORY

The analogy between the problem of inference, in the theoretical case, and the problem of acting for reasons, in the practical case, is instructive in a variety of ways, and for partisans of both debates. I have attempted to argue here that the fact that the non-standard, Anscombean story of action is in striking ways analogous to novel solutions currently being developed to the problem of inference—I have argued that this analogy is evidence in favor

---

<sup>39</sup>Or so I argue in the third essay.

of the Anscombean story: it provides us with a reason to pursue the Anscombean theory and put the Standard Story to rest.

Seeing this, of course, involves recognizing the fact that the idea of acting for reasons is under threat. This is, I think, another thing that the analogy affords us: whereas these issues arise across a wide variety of debates in action-theory, they are focused, at the moment, very narrowly in the theoretical case. The problem of inference, and the dialectic between defenders of the Taking Condition and proponents of Dispositionalism that has arisen in response to it, cleanly articulates the state of the problem of (one kind of) belief for reasons. Action theorists would do well, I think, to recognize that a similar—and similarly central—problem haunts them as well. And it is not a specter that should be ignored.

Both of these lessons are taught by the theoretical to the practical. But I think that the analogy also provides a way for the practical to give back, so to speak, to the theoretical. As I have showed, a new kind of voice is emerging in the inference literature: one that, in various ways, rejects the framework for thinking about belief and inference that has, it claims, caused the problem. The framework this new voice favors is analogous to a framework some theorists have already adopted for thinking about the nature of intentional action. And indeed, action-theorists have adopted their non-standard framework in response to a number of issues that arise across action-theory; and they have developed it with a similarly wide focus. The non-standard story, then, has been richly worked-out in the practical case; and so it seems not unlikely that philosophers of theoretical inference might have something to learn from the practical philosophers. Could, for example, the kind of Wittgensteinian emphasis given to the practical account—the idea that calling an action intentional makes reference to a form of description of events—be given to the theoretical account as well? Pursuing this question is, of course, outside of the scope of the present essay—but it does not seem to be without its own promise.

I hope, at any rate, that it is useful to recognize the homology between debates about inference and debates about intentional action. Doing so appears to bring into focus just exactly what these many debates are really about: the nature of a rational process as such, or the place of rationality and reasons in human life. And conceiving of them this way reveals the problem of inference and the problem of action to be among the most pressing problems

facing philosophers today.

## F. CONCLUSION

I have argued that novel alternatives to traditional accounts of inference bear an analogy with a certain conception of intentional action—the Anscombean story; that the novel accounts aim to solve the problem of inference in a new and promising way; and so that the Anscombean account of action is likely our best answer to the similar problem of acting for reasons. But I suggested, too, that the Anscombean account is not only best positioned to offer an account of how acting is a rational process; it furthermore seems to offer solutions to problems across ethics and metaethics. In the next essay, I pursue this thought, offering an Anscombean account of moral worth.

## IV. THE MORAL WORTH OF INTENTIONAL ACTIONS

### A. INTRODUCTION

Philosophers distinguish between two moral properties an action might have: its moral desirability, rightness, or goodness, on the one hand; and its moral worth on the other. The difference between moral desirability and moral worth is, of course, familiar from pre-philosophical moral life. Thus, suppose two agents donate equal sums to charity; one does so because it is good for her political career, the other because there are children dying of starvation who need her aid. Intuitively, both of these actions are equally morally desirable: it is a good thing that both donations should be made, that the charity should be that much richer, and that the suffering of so many children should be relieved. But there is, just as intuitively, some moral difference between the two donations: in the language of philosophers, the donation for the sake of the children has genuine moral worth; the donation for the sake of political gain lacks it.

An action is morally desirable if (and to the extent that) it is recommended (or promoted or required) by the true moral theory—whatever that may be. Or, in different terms, moral desirability is supposed to be whatever property makes an action morally good. But moral worth is supposed to be an additional property morally desirable actions may have. To explain this additional property, theorists say what is, in a way, obvious: an action can have moral worth only if it is done *for the right reasons*. The politician donates for the wrong reasons; therefore, her action, though good, lacks moral worth. The agent who is moved by the suffering of others, and wants to help alleviate that suffering, donates for the right reasons; therefore her action can have moral worth. I will refer to this account in its general form as the Right Reasons Thesis: a morally desirable action has moral worth only if it is

performed for the right reasons.

Now, theorists disagree about how to refine the Thesis in order to account for the range of subtle judgments we make about moral worth; and they disagree, too, about just what the right reasons *are*—whether they involve the concept of *rightness* explicitly, or whether they involve more determinate moral properties.<sup>1</sup> But nearly all theorists agree that some version of the Right Reasons Thesis is the correct account of moral worth.<sup>2</sup>

In this paper, I will reject the approach that has motivated the Right Reasons Thesis and propose a new approach to the problem of moral worth. In particular, I will suggest that moral desirability and moral worth are not two different properties that one and the same action may have; on the contrary, they are distinguished precisely by what they are properties of. Moral desirability is a property of action-types; moral worth, a property of action-tokens. Token actions are intentional under many descriptions—that is, they instantiate multiple types. So, I will argue, to judge the moral worth of an action we must take into account the moral value of the *many* types that the action instantiates.

The plan of the paper is as follows. I begin by briefly presenting three versions of the Right Reasons Thesis that have been proposed in the literature (§2). I argue that each existing theory is unsatisfying in some way, thereby motivating my alternative account. Furthermore, this brief survey should give us a sense of the kinds of judgments a satisfactory theory of moral worth must be able to explain. I then present my own account (§3) and demonstrate how it can accommodate all the judgments previously canvassed (§4). I go on to demonstrate another of its principal virtues: that it provides a subtle and, to my mind, correct verdict on the hotly contested debate about the moral merit of motivation *de dicto* and motivation *de re* (§5). I conclude by reviewing one final, theoretical virtue of my account

---

<sup>1</sup>A debate I will return to in §5.

<sup>2</sup>Among them, most of whom I will consider in some depth in the following, are: [Arpaly \[2002\]](#), [Markovits \[2010\]](#), [Sliwa \[2016\]](#), [Johnson King \[2018\]](#), [Stratton-Lake \[2000\]](#), and [Isserow \[forthcoming\]](#). But the Thesis is not at all new: Immanuel Kant—from whom we have inherited the language of “moral worth”—articulated a variant of it when he claimed in the *Groundwork* that only actions performed from the motive of duty have moral worth. Aristotle, too, appears to give a variant of the thesis when he claims that a virtuous action is only preformed virtuously when it is chosen “for itself.” See, for example, *NE* 1105a32. On what Aristotle means when he writes that the virtuous agent must choose the good action “for itself,” see [Whiting \[2002\]](#). Her interpretation of Aristotle’s claim—according to which choosing an action for itself involves choosing it for the qualities that make it the action that it is—brings Aristotle’s account in close proximity to contemporary Right Reasons accounts; but her insistence that this is the right way to understand how we might choose and *action for itself* brings it in even closer proximity to my own.

(§6).

## B. THREE RIGHT REASONS THEORIES

In this section I will briefly examine three Right Reasons theses. For each thesis I will argue that there is some case or kind of case that it appears to give the wrong verdict about: an action about which our intuitive judgments and the pronouncements of the theory come apart. I do not intend for my criticisms here to be decisive, however. I want, instead, to suggest that there is reason to doubt the existing theories of moral worth, in order to motivate the search for an alternative. Furthermore, by surveying the kinds of judgments that these theories have trouble explaining, we will have a battery of cases against which to test my own.

### 1. Nomy Arpaly on Moral Worth

Nomy Arpaly, who is largely responsible for bringing the debate about moral worth back into vogue, calls her version of the Right Reasons Thesis “Praiseworthiness as Responsiveness to Moral Reasons.” It runs as follows:

For an agent to be morally praiseworthy for doing the right thing is for her to have done the right thing for the relevant moral reasons—that is, the reasons for which she acts are identical to the reasons for which the action is right (the *right reasons* clause); and an agent is more praiseworthy, other things being equal, the deeper the moral concern that has led to her action (the *concern* clause). Moral concern is to be understood as concern for what is in fact morally relevant and not as concern for what the agent takes to be morally relevant.<sup>3</sup>

Arpaly’s account, as its name suggests, aims to articulate what it is for an agent to be more or less responsive to moral reasons—and it is meant to grant moral credit to those who

---

<sup>3</sup>[Arpaly, 2002, 84]. It is worth noting here that this formulation in terms of *the* right action might seem to preclude from having moral worth any actions that are not singularly called for but are still what we might call ‘morally good’. Arguably, this is already a bad result. But I am inclined to give Arpaly, as well as Markovits, who makes a similar error, the benefit of the doubt and interpret them as meaning, roughly, that the reasons for which the agent acts are the same as the reasons that the action is good. Arpaly’s examples certainly support this reading.

have more moral concern. By ‘concern,’ Arpaly means a kind of desire; and by ‘desire,’ she has in mind, roughly, the ground of a number of different dispositions—most importantly, in this case, the disposition to act.<sup>4</sup> We can measure the depth of an agent’s concern by examining how stable her disposition is. So, if the agent *would* have acted differently given minor changes in her circumstances, her disposition to act is unstable; if she *would* have acted in the same way under almost all circumstances, she shows herself to have deep moral concern.<sup>5</sup>

Arpaly’s account nicely captures our sense that those who do moral deeds in a perfunctory manner deserve less moral credit than those who seem focused on and affected by moral questions. But the counterfactual stability of an agent’s disposition to act does not appear to be a completely reliable guide to the resulting action’s moral worth. For instance, suppose two agents perform heroic ocean-rescues. One of the agents has an intense phobia of jellyfish—had there been jellyfish present, as there often are, she would not have been able to perform the rescue. Does this counterfactual render her rescue less praiseworthy than the phobia-free agent’s? I do not think that it does—yet on Arpaly’s account it seems to. This is, I think, the wrong result.

Note, too, that we need not rely on a single defeater for the disposition, like the jellyfish-phobia. Perhaps an agent is subject to a set of disabilities, all of which severely limit her ability to act in a wide range of circumstances. Nevertheless, such an agent’s action, when she finds herself in a rare position to act, should not be accorded less moral worth. Second, we can make the same point without relying on counterfactual instability based on pathologies like phobias. Perhaps it is out of deep moral concern that I avoid texting while driving—indeed, the disposition is so stable that I never have been found to text while driving, and will never be so found. (Arpaly has other measures for having ‘deep concern’ or the right intrinsic desire, which could be easily met here: I feel distressed and upset when I see people texting while driving; I notice all kinds of subtle swerves and mistakes by drivers on the road and wonder whether they are texting; etc.) And indeed, I avoid texting while driving because

---

<sup>4</sup>See [Arpaly \[2014\]](#) for a helpful discussion of desire.

<sup>5</sup>Of course, in order to be thought to have a disposition at all, there must be some counterfactual stability to the agent’s behavior; still, an agent might be thought to have a disposition to eat when hungry even if she jumps up from her plate at the slightest noise from the baby. This kind of instability does not seem to threaten our disposition-ascription.

I take seriously my obligation not to endanger other drivers on the road, out of concern for their well-being. Is this a *highly* morally praiseworthy action? I should think not; yet on Arpaly’s account, it seems to be. In fact, it appears to be more morally praiseworthy than my act of saving a drowning child, so long as it’s true that I would not have gone out into the waves had a lifeguard been on the premises.<sup>6</sup>

## 2. Julia Markovits on Moral Worth

Julia Markovits’s formulation of the Right Reasons Thesis, which she calls the Coincident Reasons Thesis, is similar to Arpaly’s Right Reasons clause:

*[M]y action is morally worthy if and only if my motivating reasons for acting coincide with the reasons morally justifying the action—that is, if and only if I perform the action I morally ought to perform, for the (normative) reasons why it morally ought to be performed.*<sup>7</sup>

This is a start, but, as Markovits recognizes, the account needs some subtilizing if it is to explain the fact that our judgments of moral worth come in degrees. Markovits’s hypothesis is that “how morally worthy a morally worthy action can be is a factor, not (as Arpaly’s discussion suggests) of whether the agent *would* have performed it in more difficult circumstances, but of how difficult the circumstances were in which the agent *actually did* perform it.”<sup>8</sup> As a measure of the difficulty, we should ask whether we, or other normal agents, would have performed that action in those circumstances. A heroic action, then—heroic actions being her stalking horse throughout this discussion—is “a right action (of some moral significance) that most of us, judging the action, would not have had the moral strength to perform, had we been in the hero’s place.”<sup>9</sup> Thus, the moral worth of an action according to Markovits will be a function of both whether the agent performs the action for the right reasons, and to what extent the agent displays extraordinary moral strength in performing it.

So, according to Markovits’s account, degrees of moral worth are credited on the basis of

---

<sup>6</sup>See Massoud [2016], Markovits [2010] and Markovits [2012] for similar criticisms.

<sup>7</sup>[Markovits, 2010, 205]

<sup>8</sup>[Markovits, 2012, 297]

<sup>9</sup>[Markovits, 2012, 297]



how extraordinary the performance of an action is—which means that, in judging the moral worth of an action, we must recur to some notion of the *ordinary*. According to Markovits, it will be the speaker’s—or judge’s—*ordinary*: the speaker refers to what is ordinary for *her* and *her kind* in judging an action’s moral worth. According to Markovits, then, moral worth is appraiser-relative.

Markovits marshals as support for this appraiser-relativism the following example: suppose a soldier runs out into enemy fire to pull a wounded comrade to safety. In response to our praise, he might say that it was “just something that he had to do”—that is, from his and his comrades’ perspective, not highly morally worthy. Markovits concludes from this that the soldier’s self-sacrifice is highly morally worthy from our point of view, but in fact not very morally worthy from the point of view of those in the military.<sup>10</sup>

But I do not think that moral worth should be understood to be relative in this way. In ordinary life, we think that the soldier is simply wrong about the worthiness of his action—we think that his action *is* genuinely extraordinary, even if he does not see it that way. Furthermore, I think we should resist a conception of moral worth according to which it may be held hostage to the local morality—no matter how depraved that morality might be. Imagine a country in which the citizens have given themselves over to a kind of porcine hedonism, to the extent that any minimally decent act toward another citizen is exceedingly rare. In such a society, helping an elderly lady having fallen on the sidewalk might be extraordinary—but surely we should not think that its citizens are in any sense correct to call such an act deeply morally worthy.

### 3. Amy Massoud on Moral Worth

Amy Massoud builds her account on state-of-the-art work on reasons done by Joshua Gert and transposed into a moral key by Douglas Portmore. Gert introduces the distinction

---

<sup>10</sup>She writes, “...different speakers will, quite properly, disagree about what counts as heroic. A striking example of this, one that reinforces the account of heroism I am defending, is the frequency with which people held up as heroes demur, and resist the label. Consider the soldier—in one sense, a member of our community—who risks his life to disarm a bomb or pull a wounded comrade out of the crossfire. Such acts will strike those of us sitting on the sidelines as heroic: we could not imagine performing them ourselves. But they may well strike the soldier as normal: any one of his comrades would have done the same. We might both be right”(297).

between the *requiring* strength and the *justifying* strength of reasons. Portmore then distinguishes the *moral requiring strength* of a reason from the *moral justifying strength* of a reason. The moral requiring strength of a reason is the extent to which that reason makes it morally impermissible to refrain from performing some act that it would otherwise be morally permissible to refrain from performing. The moral justifying strength of a reason is the extent to which that reason makes it morally permissible to perform actions that it would otherwise be morally impermissible to perform.<sup>11</sup> But for our purposes here, the point can be made somewhat more simply. There are two kinds of reasons central to Massoud’s account. First, there are moral reasons that require us (to a greater or lesser extent<sup>12</sup>) to perform actions; second, there are reasons that justify us in not performing those required actions. So, for example, the fact that I promised requires me to mow my neighbor’s lawn; but the fact that I am feeling under the weather justifies me in breaking that promise.<sup>13</sup>

For Massoud, the degree of an action’s moral worth is a function of: 1) the percentage of overlap between the agent’s motivating reasons and the moral (requiring) reasons to perform the action,<sup>14</sup> and 2) the weight of the (justifying) reasons that pull the agent in the direction of refraining from performing the action. Hence, if I have good reason not to perform a moral action, but I perform it anyway, my action has more moral worth than if I had no reason not to perform the action. The view is thus a kind of ‘self-sacrifice view:’ the moral worth of an action is increased given the weight of the reasons that the agent overcomes in performing the action—that is, by the degree to which the agent makes a sacrifice in performing it.<sup>15</sup> But Massoud’s account avoids the implausible results of some self-sacrifice views by ensuring that the reasons that pull the agent in the direction of refraining from performing the action are genuinely justifying—they are good reasons. Similarly, her account does not require that the agent be overly vexed or disturbed by the presence of such reasons, only that she be aware of them. Still, in the end, the degree of an action’s moral worth has to do with how

---

<sup>11</sup>For more on this way of classifying reasons, see Gert [2007] and Portmore [2008].

<sup>12</sup>The reasons involved in Massoud’s account come in different strengths, which are measured, more or less, by a consideration of what reasons the reason-in-question would overcome and be overcome by, counterfactually.

<sup>13</sup>Perhaps. [Massoud, 2016, 699].

<sup>14</sup>This Massoud calls the “Overlap Thesis,” and it represents her Right Reasons Clause.

<sup>15</sup>The sacrifice need not be egoistic—she might have to sacrifice, say, her daughter’s delight to perform a moral deed; nonetheless it seems natural to call this a kind of sacrifice.

much is legitimately overcome by the agent in performing the action.

I think we should find this emphasis on self-sacrifice, even thus construed, a bit dubious. Despite Massoud's best intentions, it seems to rule out at least one understanding of the genuinely virtuous action as morally worthy. For, on Massoud's account—and she takes this to be a merit—if there are *no* justifying reasons pulling the agent to refrain from performing the action, then the value of the moral reasons for which she performs the action is multiplied by zero: such an action has *no* moral worth.

Now, as Massoud notes, a reason to refrain from performing some morally good act is justifying for an agent only when the agent has epistemic access to the reason *qua* reason and moreover has sufficient reason to believe that the act it recommends would be best. But it seems to me that we can imagine an agent who is so virtuous that she will not even recognize justifying reasons *as* reasons when there is a morally requiring reason before her: she is so absorbed in administering aid after a hurricane that the fact that her house, too, is flooded strikes her as *no reason at all* to stop what she is doing. If nothing will so much as *count* as a morally justifying reason for this virtuous agent when there is a morally requiring reason on the table, then her virtuous action will always be less morally worthy than merely continent action—indeed, it will have no moral worth. This seems like the wrong result.

This is the picture of virtue put forward by virtue theorists like John McDowell (and, arguably, Aristotle). According to McDowell, facts that might serve as morally justifying reasons for merely continent agents will be *silenced*, not just overridden, for the virtuous agent.<sup>16</sup> Massoud attempts to address this issue when she asks whether an agent must “feel the pull” of her morally justifying reasons. If to “feel the pull” of countervailing reasons means to hesitate or to act without complete resolve, she writes, the answer is “no.” All that is required is “awareness or acknowledgment of the relevant self-regarding reasons,” and such awareness “need not entail that an agent acts hesitantly or without resolve.” (709) But Massoud here seems only to be countenancing what we could call *motivational* silencing: the agent in possession of a morally justifying reason need not be *motivated* at all in the direction of that reason—but, according to Massoud, she can still have it. But when McDowell, for instance, writes of the silencing of reasons, he means both *motivational* silencing and *rational*

---

<sup>16</sup>See McDowell [1998c], McDowell [1998b], and McDowell [1998a].

silencing: <sup>17</sup> “Here and now, [the consideration] does not count for him as any reason for acting...”<sup>18</sup> This, it seems, would amount to an explicit denial of the second condition of “awareness” that Massoud adopts from Markovits: that the agent must believe that the fact *is a reason*. Virtue theorists like McDowell want at least to allow for the possibility of this kind of silencing, and would presumably claim that actions performed in the presence of such silencing, as the only genuinely virtuous actions, are greatly, if not maximally, morally worthy. Indeed, this doctrine is integral to what we could call the Aristotelian doctrine of acting for the right reasons. But if there are *no* morally justifying reasons for the agent to refrain from acting, then, according to Massoud’s multiplicative account, the action has *no* moral worth. So Massoud is forced to rule out a certain conception of virtuous action as morally worthy after all.

Each of the existing Right Reasons theories, then, generates some counterintuitive results. This hardly shows that the theories are indefensible, of course; but it nevertheless opens the door for an alternative theory of moral worth. In the next section, I propose such a theory, one that eschews the Right Reasons framework altogether; in §4 I show how my account can avoid the problems I have raised in this section for existing accounts.

### C. ACTING WELL AS A DOCTRINE OF MORAL WORTH

At the beginning of this paper, I introduced the distinction between moral desirability and moral worth: moral desirability is supposed to be whatever property makes an action morally good—an action’s tendency to increase happiness, for example—whereas moral worth is supposed to be an additional property that a morally desirable action has if it is performed for the right reasons. Moral desirability and moral worth, then, are supposed by theorists to be two distinct properties that one and the same action may have.

To begin, I want to suggest that moral desirability and moral worth are not two different properties that one and the same action may have. On the contrary, moral desirability and

---

<sup>17</sup>See [Seidman, 2005, 68-77].

<sup>18</sup>[McDowell, 1998a, 91].

moral worth are distinguished by what they are properties *of*. Moral desirability is a property of action-types; moral worth, a property of action-tokens.<sup>19</sup>

What does it mean, first, to say that moral desirability is a property of action-types? Here are the kinds of claims we make about moral desirability: “Donating money to Oxfam is morally desirable because it helps relieve suffering;” “Keeping your promises is morally desirable because it is unjust to break your promises;” or “Telling the truth is morally desirable because it insults the autonomy of persons to lie to them.” The subjects in these sentences—that is, what we ascribe the property of being morally desirable *to*—are *donating money to Oxfam*, *keeping your promises*, and *telling the truth*. None of those denotes a particular event. All of them are general: they are action-types.<sup>20</sup>

An action-token, on the other hand, is a concrete particular: it is a dated, unrepeatable event. My keeping my promise to mow your lawn yesterday is a token action—it is something that I was doing, and that I did.<sup>21</sup>

Now an action-token may be morally desirable in virtue of instantiating a particular type, in the way that my action yesterday was morally desirable in virtue of its being an instance of *promise-keeping*. But virtually the first point recognized by theorists taking up the question of action in the 20th century was that an action *qua* concrete particular instantiates many different types—or, as the point was originally put: actions falls under multiple descriptions.<sup>22</sup> Donald Davidson makes the point vividly when he writes, “I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home. Here I need not have done four things, but only one, of which four descriptions have been given.”<sup>23</sup> There is one action—one unrepeatable event of

---

<sup>19</sup>The claim that “moral desirability is a property of action-types” can equally well be put, “moral desirability is a property of actions in virtue of instantiating a certain type.” I take these expressions to be interchangeable because I take the following expressions also to be interchangeable: “this type of action is desirable” and “actions of this type are desirable.”

<sup>20</sup>The claim that moral desirability is a property of types of actions is common in the history of ethics, but not undisputed. For a clear articulation of the thought, see Ross [1930], who writes, for example, of the “intrinsic rightness of a certain type of act” (47) and, interchangeably, of “*prima facie* duty’ ... as a brief way of referring to the characteristic ... which an act has, in virtue of being of a certain kind (e.g. the keeping of a promise)...” (19). For a dissenting voice, see Dancy [2004].

<sup>21</sup>Though it need not be the case that I did not mow your lawn—there might have been a token action of lawn-mowing yesterday without the lawn, finally, being mowed—if I was, say, interrupted by a car crash in front of your house. See Thompson [2008].

<sup>22</sup>See Davidson [1980], Anscombe [1957], and [Hornsby, 1980, 4], for examples.

<sup>23</sup>[Davidson, 1980, 4]. See also, of course, Anscombe [1957], from which Davidson’s account drew much

my finger moving upward—but it instantiates multiple types.<sup>24</sup>

Thus, just as an action may be morally desirable in virtue of being an instance of one type; so, if it instantiates multiple types, it seems that it may be morally desirable to some extent in virtue of instantiating one type, and morally desirable to another extent in virtue of instantiating another type. Thus we might say that my action was very morally desirable *qua* an instance of promise-keeping; but not morally desirable (or undesirable!) *qua* an instance of lawn-mowing.<sup>25</sup>

If someone inquires about the moral desirability of an action, then, we will have to ask, “under what description do you want to know if the action is morally desirable?” And if they reply that they want to know about the moral desirability of the action *simpliciter*—the action itself—then it seems that, in answering, we would have to factor in the desirability of the action insofar as it instantiates *each* of the types that it instantiates. In doing so, I claim, we are no longer talking about moral desirability—we are talking about moral worth. The question of moral worth is the question of the goodness of the action *in concreto*, and thus the goodness of the token in virtue of instantiating *multiple* types.

Of course, not every type that an action instantiates is relevant to the question of moral worth. When we ask after the moral worth of an action, I think it is clear, we are asking after the moral worth of an *intentional action*. In a world in which all actions were somnambulistic, say, it does not seem as though an opposition between moral worth and moral desirability could get a grip. Thus, when judging the moral worth of an action, we should take into account the moral desirability of the types, instantiating which, the action is intentional. As theorists have noted, an action may be intentional under multiple descriptions—or, in our terminology, an action may be intentional insofar as it instantiates *multiple* types. Hence Davidson’s action is intentional *qua* switch-flipping, and intentional *qua* light turning-on; but not intentional *qua* prowler-alerting. Thus, in judging the moral worth of Davidson’s action,

---

inspiration.

<sup>24</sup>The debate about action-identity in the 1970s rested largely on the attempt to understand the relation of the *action* to its descriptions. See, for example, Thomson [1971b], Thomson [1971a], Goldman [1971], Davidson [1980], and Anscombe [1979].

<sup>25</sup>Ross [1930] puts the point in a similar way: “...if we ask ourselves whether it is *qua* the packing and posting of a book, or *qua* the securing of my friend’s getting what I have promised to return to him, that my action is right, it is clear that it is in the second capacity that it is right; and in this capacity, the only capacity in which it is right, it is right by its own nature...” (44).

we would take into account the moral desirability of *switch-flipping* and *light turning-on*, but not *prowler-alerting*.

So here is my proposal: moral desirability—which we will call from now on, simply *goodness*—is a property of action-types, or actions insofar as they instantiate a single type. Moral worth is a property of actions, token particulars, insofar as they are intentional. And they may be intentional under a variety of descriptions—or types. Thus, in judging the moral worth of an action, we must take into account the goodness of the action in virtue of all of the types it instantiates insofar as those types describe the action as intentional.<sup>26</sup>

Let us look at a few examples of what I have in mind. The simplest cases are cases of ulterior motives. So take, again, two agents donating money to Oxfam. One does this in order to convince her constituents of something that is false; the other out of concern for others. Both actions seem to get the positive value of the action-type *giving money to charity*; but the first is marred by also attracting the description *misleading the voters*; and indeed, we can recognize that it is an instance of the more complex type, *misleading the voters by shamming virtue*. The donation motivated by concern, on the other hand, instantiates neither of those descriptions. Hence the former action is less morally worthy than the latter.<sup>27</sup>

What about cases in which the ulterior motive is not itself bad—for example, Kant’s grocer? He sets fair prices in order to make a profit: but it does not seem right to say that *making a profit* is a type of *bad* action.<sup>28</sup> It appears to make a difference, then, whether the good action is subordinate to, or in the service of, a bad or even indifferent one: we could describe this agent’s action by saying that he is *setting fair prices in order to make a profit*, or *making a profit by setting fair prices*. In this context, our judgment of the goodness of *setting fair prices* is tempered.<sup>29</sup>

---

<sup>26</sup>See also Makin [2017], which makes a similar point about judgments of the permissibility of actions according to the Doctrine of Double Effect.

<sup>27</sup>As should be obvious, many action-descriptions are morally neutral, *modulo* certain things about the contexts in which they appear. So, for example, *signing her name on this line*, a possible part of a donation to charity, strikes me as neutral, and so neither adds nor subtracts worth from the action. In general, I will not concern myself with the myriad neutral action-descriptions under which most actions fall, unless they come up explicitly.

<sup>28</sup>Though I take it that it is not accidental that there is, in fact, a perspective from which profit-making is a dubious thing to do.

<sup>29</sup>Interestingly, I think our intuitive judgments about actions described teleologically this way will distin-

Note that this tempering will not occur if two types appear side-by-side, as it were—if neither is subordinate to the other. We could imagine that the grocer sets his prices because those prices are fair, but also, separately, because those prices will allow him to make a profit. In other words, the motivation of Duty may be enough to motivate the grocer to set his prices fairly; but if setting his prices this way were not morally required, the fact that doing so would profit his business would also be enough to motivate him to do it.<sup>30</sup> I think our intuition here, perhaps diverging from Kant’s, is that his action is no less morally worthy because he is conscious of his role as a businessman. And this is borne out in the descriptions: he is not *setting fair prices in order to make a profit*, even though he is both *setting fair prices* and *making a profit* intentionally. But a neutral or indifferent action-type does not take away from a positively good one.

Next we must consider the paradigm case in these debates, the case of inadvertent virtue: Huck Finn. Huck is supposed to present a puzzle for moral theorists because he takes himself to be doing wrong when he lets Jim escape—he is acting against his moral conscience—but his action is both good and appears to be praiseworthy. Why can’t Huck bring himself to turn Jim in? In the scene in which his resolve to turn Jim in breaks down, Jim first appeals to Huck as a friend: “you’s de bes’ fren’ Jim’s ever had; en you’s de only fren’ old Jim’s got now;”<sup>31</sup> then, Jim appeals to a promise Huck has made: “de ole true Huck; de on’y white genlman dat ever kep’ his promise to ole Jim.”<sup>32</sup> The effect of these proclamations is to make Huck decide not to turn Jim in: so his failure to turn Jim in seems to instantiate the type, *helping a friend*, and indeed, *keeping a promise*. Huck believes, of course, that he is also *aiding in theft* and *wronging Miss Watson*. But we know better: you cannot steal what is not property, as a man cannot be, and likewise you cannot wrong someone by freeing her slave. So Huck’s action is neither a *stealing* nor a *wronging*. Nonetheless, I think we can say

---

guish between instrumental means and constitutive means. On the present view, one should get complete moral credit for the constitutive means to one’s ends, but not for the instrumental means. In a sense, if I am (say) *setting fair prices* in order to *make a profit*, then my commitment to performing an action under the description *setting fair prices* is highly contingent—it might be thought that I should more perspicuously describe my action simply as *taking the means (whatever they may be) toward making a profit*.

<sup>30</sup>There is nothing to prevent agents from acting for more than one reason—as Julia Tanney writes, “If someone tells you I bought my house because it was affordable, it has a big garden, and a beautiful view it would be strange to demand the reason for which I acted” ([Tanney, 1995, 108]).

<sup>31</sup>[Twain, 2005, 108].

<sup>32</sup>[Twain, 2005, 109].



that he is *doing something he believes to be wrong*—that is an accurate description of his action, and it is one that he would accept, given the crisis of conscience he is having. And in general, *doing something one believes to be wrong* is bad—but not terribly so, until we know more. So Huck’s action is a case of *helping a friend*, *keeping a promise*, and *doing something he believes is wrong*. The goodness of the first two types clearly outweighs the badness of the third.

The moral worth of Huck Finn’s action does not come away unscathed by his incorrect moral beliefs, however. The fact that he thinks what he is doing is wrong does genuinely take away from the goodness of his action, because *doing something one believes is wrong* is, I think, a bad type of action. This is reflected in our ambivalence in general in cases where the agent sincerely thinks she is doing wrong but does it anyway: on the whole, we would hesitate to encourage that kind of behavior. Still, overall we think Huck acted well—the verdict of my account.

Consider the opposite case: the bad person who believes that he does well—the murderer, say, who believes that *this* murder is necessary to liberate the human race from the bonds of capitalist greed. His action instantiates the type, *murder*, of course; and, supposing he is failing to liberate anyone with his action, we might think that that is all. But I think we should also say that his action instantiates the type, *attempting to do some good*. This does not, of course, imply that he succeeds in his further aim; but it does seem to matter that he has it. Obviously this action is wrong; but it is somewhat better, given its fuller description, than a senseless murder, or one performed out of a purely evil motive.

One might worry that I am taking for granted the most important concept: goodness—and so not really giving an account at all. But it is worth noting in this context that *all* partisans to this debate take moral desirability for granted to some extent. Goodness will ultimately, I suppose, be elucidated in terms of the correct moral theory. But the current debate takes an intuitive conception of what is good and what bad for granted. We easily call *donations to Oxfam* morally desirable and *being rude to colleagues* not; I wish to rely only on such judgments in speaking about what is good.

It is also worth noting that I rely on the fact that goodness, in this sense—as it applies to action-types—comes in degrees. But here, too, I say: all partisans to the debate took

for granted that moral desirability, whatever it may be, comes in degrees. And I think that this is no more mysterious than the fact that we call some actions morally desirable and others not. It is highly plausible, intuitively, that *saving children* is very good, while *making a friend laugh* is less good, but still good. Once again, that is all I need to rely on in developing my account.

There are further worries we might have about implementing my proposal, of course. But generally, this can be said in favor of it: it takes no more for granted than do Right Reasons theorists of the more usual stripe. They too take it that we can talk straightforwardly about the reasons for which an agent performs an action; they too talk about the moral desirability or rightness of actions without further ado. My account takes only these two categories for granted, and produces an account of moral worth in their terms.

#### D. PROBLEM CASES

In §2, I described cases and judgments on those cases that, I claimed, existing Right Reasons theories have trouble accommodating. In this section, I want briefly to return to those cases and demonstrate that my account can accommodate them straightforwardly.

The case that made trouble for Arpaly's account was that of the phobic agent who performs a heroic ocean-rescue in circumstances where, counterfactually, it was extremely unlikely that she should perform that rescue. According to Arpaly's account, where the counterfactual stability of the agent's disposition to act affects the degree of moral worth that we accord to the action she performs, this action is less morally worthy than the same action performed by an agent free of phobias—that is, whose disposition to act is quite stable. But this, I claimed, seems unfair: both rescues should strike us as equally heroic and thus equally morally worthy. According to my account, the only question we must ask is: what did the agent do? Each agent, it seems, *rushed into the waves* in order to *save the drowning child*—and since the latter is clearly very good, both agent's actions are equally morally worthy.

Our objection to Markovits relied on the case of a soldier saving his comrades in battle

and subsequently demurring, “I was only doing what I had to do.” According to Markovits, since moral worth is appraiser-relative, the soldier’s act may have some degree of moral worth from our perspective, but less from his own. But, I claimed, we should rather think that, the soldier’s protestations aside, his act is *in fact* highly morally worthy. On my account, this soldier’s action instantiates types like, *braving enemy fire* and *saving a comrade*. Altogether, it is simply a highly morally worthy action—the fact that he felt it was his duty does nothing to take away from that. Again, I think this is the correct result.

Finally, what about the virtuous agent who fails so much as to recognize reasons to refrain from doing the good deed before her? Recall that on Massoud’s account, if an agent recognizes *no* reasons to refrain from acting, then the action she goes on to perform, no matter how good in itself, has no moral worth. According to my account, however, the action of such an agent instantiates only the right types: an act of kindness will be, say, a *helping a person*, without the imposition of ulterior motives. Such an action is exactly as good as it seems: the goodness of that description gives the full account. The fact that there are no countervailing desires does not obliterate the goodness of the action: it leaves it untouched. Thus the good actions of the perfectly virtuous are indeed morally worthy. Once again, I think, this is a welcome result.<sup>33</sup>

## E. MOTIVATION *DE DICTO* AND MOTIVATION *DE RE*

I have presented what I hope is a plausible account of moral worth, and have shown how it can accommodate a wide variety of intuitions about cases. My account has a further benefit, however: it provides what I think is a subtle solution to the tricky debate over the relative moral merits of actions motivated by rightness *de dicto* and actions motivated by rightness

---

<sup>33</sup>How does my view deal with the case of the continent agent—is his action more morally worthy than the virtuous agent’s, because he overcomes countervailing desires? It is not entirely clear: it depends on whether we think that *overcoming a bad desire* is something that the agent is *doing*, and also whether or not we think that that *overcoming* is positively good. Regardless, it is not obvious that the difference between virtue and continence must reveal itself in the moral worth of such agents’ actions. Furthermore, intuitions on the difference between continent action and virtuous action diverge—Aristotelians like McDowell take virtuous action to be best of all, and continent action defective; while theorists like Massoud and Markovits appear to hold continence in higher esteem.

*de re*.

Actions motivated by rightness *de dicto* are those the performance of which is motivated by the fact that *this action is right*; the agent of such actions is supposed to have some thought like, “What’s the *right thing* to do here? I’ll do that.” Actions motivated by rightness *de re*, on the other hand, are supposed to be actions motivated by the features of the action that *make* it right—whatever those features turn out to be. If, for example, some kind of Utilitarianism were the true moral theory, we would say that an action motivated by rightness *de re* was one motivated by the fact that the action maximizes utility.

In the debate about moral motivation, some philosophers claim that motivation by rightness *de dicto* is either dubious or defective, and so actions motivated by rightness *de dicto* are not morally worthy. Only actions motivated by rightness *de re* are morally worthy according to these philosophers. Other philosophers, however, claim that motivation by rightness *de dicto* is either innocent or, in fact, paradigmatically good, so that actions motivated by rightness *de dicto* are indeed morally worthy—at least as worthy as actions motivated by rightness *de re*.

So, on one side of the debate, we find Right Reasons theorists like Arpaly and Markovits, who claim that actions motivated by rightness *de dicto* are less morally worthy than those motivated by rightness *de re*—or perhaps not morally worthy at all. Intuitively, these theorists are trying to rule out cases like that, say, of a fascist with sincere faith in the rightness of the regime. He may believe, for example, that *killing this dissenter’s family is the right thing to do*—but that does not mean that his action has moral worth. So, on the one hand, these theorists do not want to have to grant moral worth to evil actions done, misguidedly, under the guise of the “right.” On the other hand, these theorists *do* want to grant moral worth to the the actions of agents like Huck Finn, who precisely do not believe that their actions are right (when this is read *de dicto*) but who appear to be sensitive to the features that make them right—that is, to rightness *de re*. So, it is thought, motivation by rightness *de dicto* is neither sufficient nor necessary for an action to have moral worth.

Michael Smith, who introduced the idea of motivation *de dicto* in *The Moral Problem*, also doubts its moral value: he claims that “good people care non-derivatively about honesty, [...] justice, equality, and the like, not just one thing: doing what they believe to be right, where

this is read *de dicto* and not *de re*. Indeed, commonsense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue.”<sup>34</sup> His qualm is with a different kind of case from those that concerned Arpaly and Markovits, but his verdict is similarly not without intuitive appeal. If, say, we had a colleague who routinely seemed unmoved when she witnessed suffering, but who would spring into action whenever the goodness of so acting was made explicit to her, I think we should find her good deeds dubious. Intuitively, we find someone who is *only* motivated by goodness *de dicto* to be not saintly but sanctimonious, or, in any case, somewhat suspicious.

On the other side of the debate, however, are those who push back against these theorists’ dismissiveness of moral motivation *de dicto*. Paulina Sliwa, for example, has defended the merit of acting on the basis of moral advice, which seems to spell trouble for accounts that only value motivation by rightness *de re*. If I follow an experienced professor’s advice on how to discipline a student because I myself do not know what it is right to do, it seems that I am doing the right thing because it is right *de dicto* and not for the reasons it is right—after all, I do not know what those are. According to Right Reasons theorists, this action should not have moral worth. But, Sliwa asks, is it really plausible that my action here has *no* moral worth? Caught in a tough spot, morally conflicted, I reach out to a mentor for advice, and follow it—what could be wrong with that?<sup>35</sup>

Zoe Johnson King also defends the value of motivation by rightness *de dicto*. She attempts to fix our attention more seriously on the phenomenon of acting for the sake of the right *de dicto* in order to suggest that our intuitions about such cases have been misled by unfairly compared cases, and that in fact our intuition about motivation *de dicto*, when that is presented fairly, shows that we take it to be every bit as worthy as motivation *de re*.<sup>36</sup> So, for example, compare two agents: the first begins her deliberation with the thought, “I want to do what’s right here” and ends it with the intention, “So I will tell my landlord that I broke it.” The second, confronted with the same situation, straightaway decides, “I will tell my landlord that I broke it.” Is it really so clear that the second agent’s action is more morally worthy than the first? There might be two ways to interpret what is going on with

---

<sup>34</sup>[Smith, 1994, 75].

<sup>35</sup>Sliwa [2015]. See also Carbonell [2013] for another defender of *de dicto* motivation.

<sup>36</sup>Johnson King [2018].

the first agent: she may hesitate because she is motivated to do what is wrong—in which case she is continent, which is a different case—or, alternatively, she may just be more explicitly concerned with what is right. If the latter is the case it seems entirely unreasonable to grant her action no moral worth—all because she is trying to be moral! Yet her desire to come clean to her landlord is derived from a *de dicto* desire to do what is right.<sup>37</sup>

It seems to me that there are indeed some subtle distinctions controlling our intuitions about the kinds of cases that get brought up in this debate. On both sides of the debate are compelling intuitions about cases that we should like to see vindicated. My account gives, I think, a fair verdict in all these cases.

Take the example that worries Arpaly: the political extremist who wants to murder Tamara but believes, not that murdering people is wrong, but that murdering *Jews* is wrong—he recognizes that Tamara is a Jew, and refrains from killing her. He is concerned with doing the right thing, but it must be the right thing *de dicto*, because he mistakenly believes that the goodness of refraining from this murder has to do with Tamara’s ethnicity. Now, Arpaly invites us to share the intuition that his omission is not morally praiseworthy; and she wants us to believe that this is so because it is motivated by a *de dicto* belief about what is right. But it seems to me that the lack of moral worth here does not rest on any fact about the quality of this agent’s moral beliefs. His action—his intentional omission—instantiates the types *refraining from killing Tamara* and *refraining from killing a Jew*. But we can already see that those do not strike us as particularly good actions—like *not driving up onto the sidewalk and killing pedestrians*, these actions are, without further explanation, simply expected.<sup>38</sup> So we have no reason to see this action as anything like considerably morally worthy—as Arpaly insists we should not.<sup>39</sup>

---

<sup>37</sup>It is a part of Arpaly’s and Markovtis’s theories that moral worth take into account only our *intrinsic* desires, and not derived desires.

<sup>38</sup>This point should not be confused with the thought that since *not killing pedestrians* is morally *required* it does not have moral worth. We already saw, above, that required actions can be highly morally worthy. The point is, rather, that in the absence of some reason to think otherwise, we do not understand what is positively *good* about *not killing pedestrians*: if your husband came home and said proudly that he had not killed any pedestrians today, it would be appropriate to respond with some worried follow-up questions.

<sup>39</sup>Suppose the agent, rather than refraining from murder, *saves* Tamara from being murdered—again, because he believes that murdering Jews is wrong. I think we should find this act morally worthy, even though it rests on the same *de dicto* belief as the above act of refraining from murder. Hence the fact that this agent’s belief about rightness is *de dicto* does not seem to be controlling the moral worth of his actions.

Markovits, for her part, considers a case in which a bad moral authority happens to give good advice: so, for example, the very same authorities who have taught Huck Finn that Jim is mere property also teach him to give to charity. If Huck goes on to donate to charity, she claims, he is motivated by goodness *de dicto*, because his moral authority's advice is not grounded in goodness *de re*. So is Huck's donation morally worthy? Markovits thinks it is not—she thinks that the fact that some often-wrong authority says to do A is not a reason morally justifying the performance of A, and so the action does not have moral worth.

But here I think we need to flesh out the case a bit more. Does Huck give to charity because he is told that it is right—that is, does he *place the dollar in the basket* merely because some authority tells him, “that is the right thing to do”? In that case his action instantiates the types, *placing the dollar in the basket* and *doing what he believes is right*, and, for our purposes, only those. The first is neutral; as for the second, as we've already established, it is good. But it is just one good description, and it strikes us as good, but not outrageously so—it is not descriptive enough to strike us as *very* good. Surely, then, we should say that Huck's action does have some moral worth: the goodness accorded to his action by being an instance of *doing what he believes is right*. But there is nothing going for it other than that.

If Huck's moral authority were to explain to him that it is kind to help those in need, kindness is a virtue, and so on, then his action would have more moral worth, given to it by the additional descriptions it would fall under: *helping those in need*, *doing a kindness*, as well as *placing the dollar in the basket* and *doing what he believes is right*. And this, too, seems intuitive to me: the more the moral nexus of his action is explained to Huck, the more moral understanding we should accord to him, and the less it will appear to us that he is simply following an order. Right Reasons theorists should themselves be happy with this result: deepened moral understanding by its nature results in an increase in the moral worth of actions performed for moral reasons.

Markovits, however, is worried about Huck following the advice of a certain moral authority—the same moral authority who claims that Jim is mere property. She thinks that it is merely accidental that Huck will perform the right action if he is following that authority's advice, and this kind of accident is what Right Reasons theories aim to elimi-

nate. But it is not so clear to me that she is right here. As we all know, Aristotle had some unsavory views on the moral status of women, for example. But I wouldn't fault a student of mine for reading Aristotle and deciding to cultivate in herself the virtue of courage, say—and I wouldn't think that, if she brings it off, she is just accidentally virtuous, or not really virtuous at all. In part, I think this is because Aristotle's views on women are importantly detachable from his conception of courage. And as there, so too here: if Huck's moral mentor's beliefs about charity are recognizable to us, then they will be suitably detachable from any views he has on the status of slaves. This leaves his positive views about kindness and so on untouched, and so the connection of his advice to the goodness of kindness will not be accidental.

Return, then, to my verdict that Huck's action has some moral worth because it falls under the description *doing what he believes is right*. Does this not welcome into the sanctum of moral worth all manner of despicable actions: the robbery, the murder, the genocide, motivated by the fervent belief *that it is right*? Take one of these actions: a pious political murder will be an instance of, say, *murdering this man, upholding the fascist regime, getting rid of the dissidents*, and also *trying to bring about the kingdom of Heaven on earth*, and therefore *trying to do the right thing*. So it has two good action-types to its name—but obviously that is vastly outweighed by the evil action-types it exemplifies. This is a bad action, overall. But I do think it is better than an action, utterly imaginary perhaps, that instantiates *murdering this man, upholding the fascist regime, getting rid of the dissidents*, and *doing the evil thing*—that is, an action done under the guise of the bad. At least our agent believes that he is doing something good—this *is* (the slightest bit) redeeming, and we should prefer to see agents trying to do what is right and getting it massively, disastrously wrong than to see agents simply aiming at evil.<sup>40</sup>

So, I claim, acting for the sake of the right *de dicto* provides the action with one good

---

<sup>40</sup>For a discussion of the possibility of this kind of agent, see [Velleman \[1992\]](#). But we must not be misled by the topic of Velleman's inquiry—so-called Guise of the Good Theories—into thinking that perhaps *all* actions, as performed under the guise of the good, attract the description *doing something good*. Even if all actions are performed under the guise of the good, that is not enough to impart the description *doing what she believes is good* to the action—just as it would be wrong for the Humean about acting to claim that every action is at the same time an intentional *satisfying a desire*. The Guise of the Good thesis, like the Humean Theory, provides a formal account of action which is pitched at a different level than that of reasons in the sense that we have taken them up here.



action-description: *doing what he believes is right*; this has positive moral worth. But this small amount of positive moral worth can be easily outweighed if the action's other qualities are bad; and it can be far over-shadowed by an action performed for more substantive moral reasons. And this seems like the right result: it is perhaps too churlish to say that agents motivated by the right *de dicto* deserve no moral credit, but we can appreciate that actions motivated *de re* will likely have much more moral worth.

The final kind of case to consider is Johnson King's: an agent who has the intrinsic desire to act rightly, and then deliberates until she has figured out what would constitute acting rightly. Here is Johnson King's description of the case: "Maryam is chairing a session at a prestigious Philosophy conference, which is notorious for getting nasty during Q&A. Maryam wants to act rightly—that is, she wants to conduct Q&A in such a manner as to meet all of her obligations not only *qua* chair but also *qua* moral agent. So she thinks carefully about what her obligations might be, planning to modify her behavior in light of her conclusions. After much soul-searching and careful thought, Maryam decides that four things matter morally in her case ..." <sup>41</sup> Johnson King invites us to share the intuition that Maryam's action is highly morally worthy: she is, Johnson King declares, a "moral saint." But since her motivations are all derived from a *de dicto* motivation to act rightly, her action has no moral worth, according to Right Reasons theories. As should be clear, my account can accommodate this case quite straightforwardly: if Maryam has gone through sufficient moral deliberation, then her actions will be instances of *allotting time fairly*, *discouraging rudeness*, etc., as well as *doing what she believes is right*. These are all good action-types, contributing to this action's positive moral worth.

Maryam's counterpart, Mario, is supposed to have all the desires that Maryam ends up with after deliberation, but have them *de re*—in him, they do not have to be derived. As Johnson King puts it, "Mario introspects and finds that he has four intrinsic motivations relevant to his circumstances..." <sup>42</sup> Motivated by these four intrinsic desires, Mario too *allots time fairly*, *discourages rudeness*, etc. But he is not thinking about doing the right thing, and so his action does not instantiate that type. On my account, then, it seems that Mario's

---

<sup>41</sup>[Johnson King, forthcoming, 8].

<sup>42</sup>[Johnson King, forthcoming, 9].

actions are slightly less morally worthy than Maryam's. And though this contradicts the intuitions of some Right Reasons theorists, it seems, at any rate, like a plausible result.

In general, though, this does not render indefensible the Right Reasons theorists' preference for moral motivation *de re*. I think the thought is this: we take agents motivated by rightness *de re* to be *sensitive* to the moral landscape in a way that agents motivated merely by rightness *de dicto* are not. And when an agent is sensitive in this way, I submit, she can naturally be understood to be in contact with *more reasons* than an agent who is insensitive. The morally sensitive agent will notice, for example, the passing frown on her interlocutor's face, as well as the strained smile that follows it, the fidgeting discomfort, etc. All of these will be reasons for her to, say, control her rising temper—her action will thereby be, for example, a *trying to ease her interlocutor's discomfort, trying to avoid making her upset, relieving the pressure on her interlocutor to act naturally when she feels upset*, and so on. Sensitivity, in other words, increases the number of intentional types one's action instantiates, thereby increasing the potential goodness of the action (and, of course, the potential badness). When the Right Reasons theorist imagines the agent motivated by rightness *de re*, she is imagining, I think, an agent who is sensitive in this way: one whose actions are refracted, as it were, through a prism.

As Johnson King imagines Mario, he is not really all that sensitive—he just happens to have a (limited) set of good intrinsic desires. So it is no surprise that his action is not more morally impressive than that of someone who is thinking explicitly about the good. But he is not, I think, very like the Huck Finn characters of Right Reasons theorists' dreams.

## F. CONCLUSION: NON-ACCIDENTALITY

I would like to conclude by considering one last, more general virtue of my account. In the background of some, and in the foreground of other, debates about moral worth there is what we can call the Non-Accidentality condition, which states: a morally worthy action cannot be an instance of someone's accidentally doing the right thing; that is, it must not be an accident that the agent performs the action that is recommended by the correct moral

theory.<sup>43</sup> This condition is highly intuitive—indeed, it appears to capture part of what we *mean* in calling some actions morally worthy and some not. A donation made for the sake of political gain is not morally worthy because it appears to be a matter of *accident* that the agent performed a *good* deed at all: it is just lucky for us, or for the starving children, that what is politically expedient also happens to be good. Genuinely morally worthy actions, we think, cannot be this way.

The Non-Accidentality condition both informs the way authors develop theories of moral worth and serves as an argumentative tool used against those theories. So, for example, Arpaly’s *moral concern* view, according to which the moral worth of an action is determined in part by the stability of the agent’s disposition to so act, can be understood as responsive to the Non-Accidentality condition: counterfactual stability seems to be just one way we can capture the idea of something’s being non-accidental. Markovits’s and Massoud’s accounts can be given a similar reading: the moral strength required of an agent, and the strength of the justifying reasons she overcomes in performing some action, both appear to be ways of measuring how surely the agent was going to act—that is, the degree to which it is *no accident* that she did.

In the debate about moral motivation *de dicto* and *de re*, on the other hand, the Non-Accidentality condition is used to prove one or the other kinds of motivation unsuited for morally worthy action. So, for example, both Sliwa and Johnson King argue that theorists who favor motivation by rightness *de re* appear to accord moral worth to accidentally right actions: since the agent is not thinking explicitly about the rightness of her act, she risks responding to features of a situation even when they are vastly outweighed by moral concerns.<sup>44</sup> But what’s sauce for the goose is sauce for the gander: Arpaly and Markovits, for example, both argue that if we praise an action that is motivated by the thought that *this*

---

<sup>43</sup>The first formulation is close to that of [Johnson King \[2018\]](#); the second, [Lord \[2017\]](#). See also [Isserow \[forthcoming\]](#) for a discussion of what she calls the “non-accidentality constraint.”

<sup>44</sup>See [Sliwa \[2016\]](#) and [Johnson King \[2018\]](#). Sliwa, for instance, is thinking about a case in which I am motivated by the fact that I can save a friend some embarrassment by driving her to the meeting for which she is about to be late, and argues that I might as well be motivated by the fact that I can save my friend some embarrassment by killing her ex-boyfriend, since at no point am I explicitly wondering about what is right to do in the situations. This example is not particularly convincing, of course: being sensitive to someone’s embarrassment does not preclude being sensitive, also, to peoples’ right to life; but it allows Sliwa to forge onward to the second horn of the dilemma she formulates for advocates of motivation by rightness *de re*.

*action is right*, then we might end up praising accidentally right actions: for example, the vegetarian diet kept by the Nazi because the Führer told him that it, among other things, was right.<sup>45</sup>

So the Non-Accidentality condition has proven a centrally important condition on a theory of moral worth. Note, then, what my account has to say about the Non-Accidentality condition: it meets it trivially—or perhaps more precisely, it makes nonsense of the condition. What my account brings into central focus is the fact that there can be no question of *accident* in the relation of an action to its reasons: an action is constituted by its reasons, in the sense that the descriptions under which it is an intentional action—that is, the intentional types it instantiates—are given by the reasons for which it is performed. So there can be no question of accident in the relation of the *right* action to the *right* reasons either. It is simply not the right action, when considered in full, unless it was done for the right reasons.

At the outer limit, this point is obvious: if, say, I turn on the light in order to read, and I thereby alert the prowler, it is “accidental” that I alerted the prowler in the sense that I didn’t *do* that at all: it was not an intentional action of mine.<sup>46</sup> And so of course I did not alert the prowler for the right reasons: I did not alert the prowler for any reasons at all. But the more careful we are in describing the actions we do perform, the clearer it will be that this point comes all the way in, so to speak: if I did not give to charity in order to provide the children with food, then it cannot be said in the full-blooded sense that I provided the children with food, even if, as a result of my donation, the children were fed. Or, to put the point most succinctly: I cannot *do anything* (intentionally) accidentally.

From this perspective, then, the only thing that might be “accidental” is that one action should share a description with another: that, for example, both an *political maneuver* and a *charitable deed* could share the description *donating to Oxfam*. But the Non-Accidentality condition is not concerned with a single action-description, even if its representation makes it appear to be. It is concerned, rather, with the performance of an action, and *that* is characterized not by a single description which it might share with other performances, but by an entire network of descriptions. Once we appreciate this fact, we can rest easy that the

---

<sup>45</sup>[Arpaly, 2002, 73 ff.] and Markovits [2010].

<sup>46</sup>Or, to use another form of expression, alerting the prowler was not a *human action*, even if it was an *act of a human being*—see Anscombe [2005].

Non-Accidentality condition will always be met—always, that is, when a fuller description of the action actually performed is given.

I have argued that we should conceive of moral worth as the goodness of an action insofar as it instantiates multiple different types; moral desirability is the goodness of individual action-types. Moral desirability can attach to actions *in abstracto* as much as to actions *in concreto*: moral theories may formulate deontic statements about action-types, and we may determine whether or not a particular action, in virtue of being of a certain type, is morally desirable. Moral worth, however, is a property of actions *in concreto*: particular performances have moral worth accorded to them, as it were, after the fact.

My account is not meant to be entirely at odds with Right Reasons theories. There is much to recommend such theories, and I hope to have retained all that there is to do so. The effect of my account is rather to refocus our philosophical sights on *actions* instead of *reasons*. The debate about moral worth has proceeded under the assumption that our actions and our reasons can be separated; I think this assumption is false. I have attempted to provide a way for us to think about actions as inseparable from reasons, by attending to the relation of our reasons to the types our actions instantiate—though of course more work is needed to defend this idea against arguments and intuitions to the contrary. Still, I hope that the idea shows its value by its fruits: if we assume that reasons and actions are more inextricably entwined than has previously been supposed, then the problems that arise when we separate reasons from actions—like the puzzles about moral worth—do not arise. And this is, I think, a good result.

## V. CONCLUSION

My topic, in this work, has been the idea of a rational process, or rational act, as such; my aim has been to urge a particular conception of this. Reason has many fruits: among them are explicitly cognitive attitudes, like belief and knowledge; other attitudes, like approval or disapprobation; and events like actions and reactions. One way to understand the contribution of reason to all of these things is by conceiving of reason as one of many causal inputs into a process with a particular result. But I think that understanding any of reason's products in this way underestimates and, thus, mischaracterizes what is distinctive about reason's contribution. If the typical conception places reason in the role of sculptor, and belief, action, and so on in the role of sculpture, then I have tried to urge a conception according to which reason is, instead, the shape of the sculpture. Reason is not external to its product, but shapes it; the product is only the particular product of reason that it is because reason structures it; we are only able to see the product for what it is by seeing reason in it.

In particular, of course, my argument has concerned understanding actions in this way. This is valuable, first, because I think it is relatively uncommon to conceive of actions at this level of abstraction. All too often the question of the production of action is taken to be, if not explicitly, a kind of psychological question, wherein we seek some mental states or others to play the role—switching metaphors—of motor for the movement of an all-too-physical substance. The most plausible mental states to play this role are those that can be given a reasons-conducive analysis, since, as everyone recognizes, actions are often performed for reasons. And so reason itself becomes mechanized and psychologized.

My first essay meant to remedy this tendency by showing the close similarity between the question of inference and the question of acting for reasons. Inference—though itself

in some danger of being psychologized—is generally recognized to inhabit the rarefied air of the Rational: for though inference is, of course, instantiated in a thinker’s psychology, it represents an intersection of psychology and logic, and so is given an account that rises to the latter topic. The similarities that the concept of action bears to the concept of inference reveal that it, too, is a topic that is importantly at the intersection of psychology—or physiology, or whatever—and reason and the normative, and so must be given a similarly formal account.

Examining actions as a means of engaging with the topic of a rational act as such seems valuable, second, because of the rich set of materials actions provide for thought. Consider my Aristotelian metaphor: the sculpture is made the particular sculpture that it is by means of its shape; so too is a rational act made the act that it is by means of reason. We make the point vivid by saying that without the shape, what could be a sculpture is nothing but a formless lump of clay; it is in this sense that it is constituted as a sculpture by its shape. Now take something like belief, which, I think, is made what it is in this way. What is supposed to correspond to the lump of clay? It is very hard to imagine some raw material in the case of belief; we might come close to it by imagining the declarations of the acutely insane, but the thought is too sad to linger on for long. Yet without some sense of the raw material, it is somewhat harder to grasp the transformative, constitutive force that reason plays *vis a vis* belief.

With actions, however, the picture is much clearer. Imagine, for example, a comic friend who trips, flailing, over a carpet—and then looks up at you and winks, revealing that the gaff was a fake. At first the physical movements appear dumb, a series of unfortunate events; but when they are revealed to be the product of reason, there is a gestalt switch, and suddenly the whole is revealed as an action. With actions, in other words, we know all too well what the clay is: mere, thoughtless physical movements; and, with these in view, we have a much clearer sense of how reason imbues the whole with an entirely different kind of being.

My second essay made use of progress made by philosophers considering the theoretical inference in order to support a particular conception of action—one that underwrites an understanding of action as constituted by reason in this way—as the way forward on the

question of acting for reasons. But ultimately, I argued, it is the practical case that is poised to lead the way in this novel approach to the question of a rational act as such.

Finally, it seems to me that illuminating the nature of actions as rational in the sense I have outlined is good not only in itself but also in virtue of its consequences. As is clear, the conception of action that I have urged has been available at least since Elizabeth Anscombe presented it in *Intention* in 1957; yet it seems to me that the insights afforded by this conception of action have not been fully assimilated in the ethical literature. When actions appear in ethical debates—as, of course, they very often do!—they are left more or less unquestioned, or treated as stable, atomic units. A more subtle understanding of the nature of actions—one suitably informed by the thought that actions are constituted by their reasons—has a lot to offer our ethical theorizing. In the third essay, I made a case study of this thought.

Much of what this idea of action has to offer has been obscured, I think, by the contemporary prominence of *reasons* in ethics and metaethics. A *reason* has lately been the normative notion *du jour*, with a number of theorists even contemplating a reduction of all other normative notions to reasons. Analyses of ethical problems in terms of reasons are, then, ubiquitous; it is to reasons that we have become accustomed to turn when addressing a metaethical problem.<sup>1</sup>

Reasons are, of course, reasons *for* something—there are no plain reasons, but only reasons that relate to some act: in the practical case, actions. But it seems to me that reasons theorists, in virtue of their confidence, first, in the power of reasons; and in virtue of their focus, second, on the reason-relation itself, or the fact that is the reason—reasons theorists tend to take for granted the other terms that enter into the reason relation. Actions are, therefore, in the background of a number of metaethical debates, but they are left unquestioned, treated as atomic, uninteresting terms, while reasons soak up the spotlight.

I have argued that actions—performed actions—cannot be understood apart from their reasons. And so it seems to me, in the same way, a mistake for reasons theorists to treat

---

<sup>1</sup>This era might be drawing to a close, with a group of philosophers turning to the normative notion of *fittingness* as able to do what reasons couldn't and, therefore, as more suited to be the fundamental normative concept. I think this movement is as doomed as the reasons fundamentalist movement was, but I will not enter into this debate here.



actions as atoms that can be taken for granted when they make mention of reasons. This thought, however, depends on something that I have not yet argued for. In particular, if I am to claim that reasons theorists make a mistake when they take this *relatum* of the reason-relation to be a simple entity, in need of no analysis itself, then I must claim not only that performed actions are internally complex, but also that actions *in prospect* are internally complex. Reasons, after all, when they are not used in the rationalization of an already performed action (“he got the wrench in order to fix the sink”), but rather are used to promote a not-yet-performed action (“you ought to call your mother, since she loves you so”)—when reasons are used to promote something to be done, they take a prospective action as their object. This is, indeed, the primary form reasons take. Thus, if I claim that reasons theorists are wrong to treat the *relatum* of reasons facts as single action-types, exhausted by a single description, then that must be because actions thought of in prospect, and not just concrete actions, are also internally complex, and thus, inseparable from reasons.

This claim is not one that I am in a position to argue for here. But it does seem to me to have intuitive appeal. To begin with, it seems to me that even when we mention actions in prospect by means of a single action-type, we mean something more complex. Suppose I find out that my wayward son has stolen twenty dollars from the local grocer; handing him the twenty dollars I found in his wallet, I say, “I know that you stole this from the grocery store; I think you ought to go right back there.” Later I find out that he went right back and, indeed, stole some more. Surely I might say “that’s not what I meant!”—and surely that might be true. Strictly speaking I made mention of a reason for him to return to the store—and return he did. But it seems clear that I meant that he return to the store as a part of returning the money; the *returning to the store* that I was urging was precisely *not* one that was a part of another theft.

Or again, suppose that I counsel a friend, “Mrs. Horner has really been down lately; it would be so kind of you to pay her a visit today;” then my neighbor offers this counsel: “Mrs. Horner has a ton of extremely valuable silver in her apartment, and she has no idea what it’s worth! I say you pay her a visit...” Once again, appearances suggest that both my neighbor and I have offered reasons for the same action: paying Mrs. Horner a visit. But surely it is intuitively correct to say that nonetheless we do not really suggest doing the

same thing.

Furthermore, sensitivity to the kinds of distinctions that these examples make obvious, I think, provides answers to some of the puzzles that arise for theories of reasons and for theorists as they go about applying reasons to problems in metaethics. That is, the thought that reasons relate not to single action-descriptions but to something more complex has potential to clear up confusions that are introduced by unexamined use in metaethics of reasons for actions. I will mention one example briefly.

Certain value-concepts seem to lend themselves readily to an analysis in terms of the responses they elicit in agents. This is especially clear in cases where the value-term has the suffix “-able,” like *desirable* and *admirable*—such terms seem to make explicit reference to themselves as the object of a capacity or attitude. This fact, perhaps, has given rise to an ambition among philosophers to analyze all value-concepts in terms of the responses they elicit in agents: to provide what are called “fitting attitude accounts” of value, *simpliciter*, as well as of more determinate value-concepts.

Philosophers interested in analyzing all normative phenomena in terms of reasons, then, analyze value-concepts, through the idea of eliciting attitudes, in terms of reasons, claiming, for example, that what is valuable is *what one has most reason to value*. Theorists can give accounts of all kinds of value-concepts on this model: we can define what is *admirable* in terms of what we have most reason to admire; what is *fearful* in terms of what we have most reason to be afraid of; and so on.

But a well-known problem arises for these kinds of accounts. Take the more concrete case of admirability. The account would state that what it is for X to be admirable is for there to be sufficient reason to admire X. For such an account to be attractive, whatever reasons there are to admire X should come from X’s own properties: since an object (or most often, a person) is admirable because of qualities that that object (or person) has, so the reasons in terms of which admirability is analyzed must also be grounded in that object’s (or person’s) qualities.

But philosophers can arrange that this is not so. Indeed, we can quite easily come up with reasons to admire X that seem to have nothing at all to do with whether or not X is admirable. If that is so, then the account in terms of reasons fails as an analysis.

This is the ‘Wrong Kind of Reason’ objection to reasons-based fitting-attitude accounts of value. The proponent of such an account must identify some principled way to distinguish the right kind of reasons from the wrong kind of reasons in order for her account to be compelling.

It seems to me that close attention paid to the intricacies of the action (or attitude) recommended by the reason is illuminating here. Take a classic example: suppose an evil demon threatens to take my child’s life if I do not admire him. This seems to supply a (very good!) reason for me to admire the demon; yet this reason appears to be of the wrong kind, and therefore should not count to show that the demon is himself admirable.

The problem gets off the ground, it seems, because we can find no principled distinction between two possible reasons to admire the demon: the fact that the demon is powerful and just may be a reason to admire him; but the fact that if you don’t admire him, he will kill your children is also a reason to admire him. The challenge is to distinguish the two in some way other than saying that one is the wrong kind and the other the right.

If reasons related not to single action-types but, implicitly, to strings of action-types, more complex wholes that bear the stamp of reasons, then I think we would have a way to distinguish these kinds of reasons. For the fact that the demon threatens to kill my child is a reason to admire him only when that admiring is a part of a larger action, namely, a ‘making sure the demon doesn’t make good on his threat’ or a ‘protecting my child.’ And in general, all the ‘wrong kind of reasons’ will have this form: they will promote the attitude only insofar as that attitude is a part of a larger act—whatever is motivated by the reason. The right kind of reasons will never be like this.

That is, of course, by no means a full treatment of the topic of reasons-based fitting-attitude accounts of value, but I hope it is provocative. If we could train ourselves to see the actions to which reasons relate as having a rich character, rather than being single action-types, then I suspect a number of problems in metaethics could find resolution.

So there seem to me to be good reasons to pursue an account of actions-in-prospect as having complex internal structure. But there will also be difficulties in doing so. The most glaring, I think, is this. The way I prefer to talk about this complex notion of action-in-prospect is by means of *strings* of action-descriptions, or -types: a reason, then, would

relate (implicitly) to a string of action-descriptions or -types, rather than a single action-description, or single action-type. But if I think that all thought of action is thought of something internally complex, then it seems that this should hold, as well, of the action-types that make up the string that characterizes the action-in-prospect. In other words, it seems that in formulating what I mean when I say that actions-in-prospect are internally complex, I make use of the very thing that I say is a product of confusion: namely, single action-descriptions.

One temptation might be to say that the action-descriptions that make up the strings that are, on my account, genuine actions—the single action-descriptions—are not, themselves, as Anscombe would say, “formally descriptions of intentional actions”—they are, rather, descriptions that are compatible with their object’s being either an action or a mere event; they are silent on that question. But this would be an undesirable compromise: if I think that reasons constitute an action as the action that it is, I should not like to recognize a true description of an action as one which is silent on the question of whether it was performed for reasons. That would seemingly undo what is done by my insistence on the very tight connection between reasons and actions, and would pose anew the question of what the connection, then, must be.

Another question to be addressed by this account would be what, exactly, holds the types or descriptions together in thought of an action-in-prospect. Performed actions provide an easy answer to this question: the fact that the action in question is truly describable by all of the descriptions, or in fact instantiates all the types—this is what makes it true to say that one and the same action is both a brick-laying and a house-building. But this expedient is, clearly, unavailable when the action has not yet been performed.

The answer that most naturally suggests itself is that the reason itself holds the string of action-descriptions together. After all, the A-D order represents an order of reasons, as well as a set of descriptions of an observed action; the D-A order just represents a process of reasoning. Perhaps the fact that I promised is a reason to mow my neighbor’s lawn; the lawn-mowing recommended is, I want to say, also a promise-keeping: the reason constitutes the action even in prospect.

This is, I think, what is going on: but if it is so, then reasons are going to be much

different from what we previously supposed. The “reason-relation” often adverted to in accounts of reasons is sometimes described as one thing’s “counting in favor of” another: P is a reason to A only if P counts in favor of A (or doing A, or A-ing). Something’s counting in favor of another is a relatively intuitive notion; but it is also one that appears to require a distinction between the two things it relates. If I am right, any mention of a reason will be already internally related to the action it is a reason for; reasons will not be independent or self-standing, but will represent, rather, just so many determinations of an action—or a belief, or another attitude. The account, then, would require that the nature of reasons be radically re-thought.

All this is work that has yet to be done. But I hope that what I have done here is suggestive. Acting for a reason is not like jerking in response to an electric shock: reasons are not one cause among many in a chain of causes that concludes in an action. Human behavior is transformed by its relation to reasons into human action; the latter has no existence apart from reasons, and cannot be known except through reasons. Keeping this idea clearly in mind—about actions, but also about beliefs, as well as other attitudes—will, I hope, free us from the apparent philosophical problems that arise when we do otherwise.

## BIBLIOGRAPHY

- Jesús H. Aguilar and Andrei A. Buckareff. *Causing Human Actions: New Perspectives on the Causal Theory of Action*. Bradford, 2010.
- G. E. M. Anscombe. *Intention*. Harvard University Press, 1957.
- G. E. M. Anscombe. Under a description. *Nouûs*, 13(2):219–233, 1979.
- G. E. M. Anscombe. St. Andrews Studies in Philosophy and Public Affairs, 2005.
- Nomy Arpaly. *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford University Press, 2002.
- Nomy Arpaly. Duty, desire and the good person: Towards a non-aristotelian account of virtue. *Philosophical Perspectives*, 28(1):59–74, 2014.
- Paul Boghossian. What is inference? *Philosophical Studies*, 169(1):1–18, 2014.
- John Broome. Practical reasoning. In José Luis Bermúdez and Alan Millar, editor, *Reason and Nature: Essays in the Theory of Rationality*, pages 85–111. Oxford University Press, 2002.
- John Broome. Comments on boghossian. *Philosophical Studies*, 169(1):19–25, 2012.
- Vanessa Carbonell. De dicto desires and morality as fetish. *Philosophical Studies*, 163(2):459–477, 2013.
- Lewis Carroll. What the tortoise said to achilles. *Mind*, 4(14):278–280, 1895.
- Jonathan Dancy. *Practical Reality*. Oxford University Press, 2000.
- Jonathan Dancy. *Ethics Without Principles*. Oxford University Press, 2004.
- Jonathan Dancy. Action, content, and inference. In P. M. S. Hacker and Hans-Johann Glock and John Hyman, editor, *Wittgenstein and Analytic Philosophy: Essays for P.M.S. Hacker*. Oxford University Press, 2009.
- Jonathan Dancy. *Practical Shape: A Theory of Practical Reasoning*. Oxford University Press, 2018.

- Donald Davidson. *Essays on Actions and Events*. Oxford University Press, 1980.
- Anton Ford, Jennifer Hornsby, and Frederick Stoutland. *Essays on Anscombe's Intention*. Harvard University Press, 2011.
- Joshua Gert. Normative strength and the balance of reasons. *Philosophical Review*, 116(4): 533–562, 2007.
- Alvin I. Goldman. The individuation of action. *Journal of Philosophy*, 68(21):761–774, 1971.
- Ulf Hlobil. Against boghossian, wright and broome on inference. *Philosophical Studies*, 167 (2):419–429, 2014.
- Ulf Hlobil. In defense of the taking condition. 2016.
- Ulf Hlobil. Inferring by attaching force. *Australasian Journal of Philosophy*, pages 1–14, forthcoming. doi: 10.1080/00048402.2018.1564060.
- Jennifer Hornsby. *Actions*. Routledge and Kegan Paul, 1980.
- Jessica Isserow. Moral worth and doing the right thing by accident. *Australasian Journal of Philosophy*, pages 1–14, forthcoming.
- Zoë A. Johnson King. Accidentally doing the right thing. *Philosophy and Phenomenological Research*, 2018.
- Zoë A. Johnson King. Praiseworthy motivations. *Noûs*, forthcoming.
- Irad Kimhi. *Thinking and Being*. Harvard University Press, 2018.
- Douglas Lavin. Three comments on joseph raz's conception of normativity. *Jurisprudence: An International Journal of Legal and Political Thought*, 2(2):329–378, 2011.
- Errol Lord. On the intellectual conditions for responsibility: Acting for the right reasons, conceptualization, and credit. *Philosophy and Phenomenological Research*, 95(2):436–464, 2017.
- Stephen Makin. Double effect as action theory. Manuscript, 2017.
- Julia Markovits. Acting for the right reasons. *Philosophical Review*, 119(2):201–242, 2010.
- Julia Markovits. Saints, heroes, sages, and villains. *Philosophical Studies*, 158(2):289–311, 2012.
- Amy Massoud. Moral worth and supererogation. *Ethics*, 126(3), 2016.
- John McDowell. Are moral requirements hypothetical imperatives? In *Mind, Value, and Reality*, pages 50–73. Harvard University Press, 1998a.

- John McDowell. The role of eudaimonia in aristotle's ethics. In *Mind, Value, and Reality*, pages 50–73. Harvard University Press, 1998b.
- John McDowell. Virtue and reason. In *Mind, Value, and Reality*, pages 50–73. Harvard University Press, 1998c.
- Conor McHugh and Jonathan Way. Against the taking condition. *Philosophical Issues*, 26(1):314–331, 2016.
- Ram Neta. What is an inference? *Philosophical Issues*, 23(1):388–407, 2013.
- Sarah K. Paul. The conclusion of practical reasoning: The shadow between idea and act. *Canadian Journal of Philosophy*, 43(3):287–302, 2013.
- Douglas Portmore. Are moral reasons morally overriding? *Ethical Theory and Moral Practice*, 11(4):369–388, 2008.
- Joseph Raz. *From Normativity to Responsibility*. OUP Oxford, 2011.
- Sebastian Rödl. The single act of combining. *Philosophy and Phenomenological Research*, 87(1):213–220, 2013. doi: 10.1111/phpr.12043.
- David Ross. *The Right and the Good*. Oxford University Press UK, 1930.
- T. M. Scanlon. *Being Realistic About Reasons*. Oxford University Press, 2014.
- Mark Schroeder. The humean theory of reasons. In Russ Shafer-Landau, editor, *Oxford Studies in Metaethics Vol. 2*, pages 195–219. Oxford University Press, 2007.
- Scott R. Sehon. Deviant causal chains and the irreducibility of teleological explanation. *Pacific Philosophical Quarterly*, 78(2):195–213, 1997. doi: 10.1111/1468-0114.00035.
- Jeffrey Seidman. Two sides of 'silencing'. *Philosophical Quarterly*, 55(218):68–77, 2005.
- Paulina Sliwa. Praise without perfection: A dilemma for right-making reasons. *American Philosophical Quarterly*, 52(2), 2015.
- Paulina Sliwa. Moral worth and moral knowledge. *Philosophy and Phenomenological Research*, 93(2):393–418, 2016.
- Michael Smith. The humean theory of motivation. *Mind*, 96(381):36–61, 1987.
- Michael Smith. *The Moral Problem*. Blackwell, 1994.
- Philip Stratton-Lake. *Kant, Duty, and Moral Worth*. Routledge, 2000.
- Julia Tanney. Why reasons may not be causes. *Mind and Language*, 10(1-2):103–126, 1995.



- Michael Thompson. *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press, 2008.
- Judith Jarvis Thomson. Individuating actions. *Journal of Philosophy*, 68(21):774–781, 1971a.
- Judith Jarvis Thomson. The time of a killing. *Journal of Philosophy*, 68(5):115–132, 1971b.
- Mark Twain. *The Adventures of Huckleberry Finn*. Dover Publications, Inc., 2005.
- Markos Valaris. Reasoning and regress. *Mind*, 123(489):101–127, 2014. doi: 10.1093/mind/fzu045.
- Markos Valaris. What reasoning might be. *Synthese*, 194(6), 2017a. doi: 10.1007/s11229-016-1034-z.
- Markos Valaris. What the tortoise has to say about diachronic rationality. *Pacific Philosophical Quarterly*, 98(S1):293–307, 2017b. doi: 10.1111/papq.12172.
- J. David Velleman. The guise of the good. *Noûs*, 26(1):3–26, 1992.
- Ralph Wedgwood. The normative force of reasoning. *Noûs*, 40(4):660–686, 2006.
- Jennifer Whiting. Eudaimonia, external results, and choosing virtuous actions for themselves. *Philosophy and Phenomenological Research*, 65(2):270–290, 2002.
- Bernard Williams. Internal and external reasons. In Kieran Setiya and Hille Paakkunainen, editors, *Internal Reasons: Contemporary Readings*. MIT Press, 1979.
- Barbara Winters. Inferring. *Philosophical Studies*, 44(2):201–220, 1983.
- Crispin Wright. Comment on paul boghossian, “what is inference”. *Philosophical Studies*, 1(1):1–11, 2012.